January 2015

A Subcontinental Reconstruction of Invasion Patterns and Processes for the Past Two Centuries

Teresa Lynn Clark
Purdue University

Graduate School Form 30
Updated 1/15/2015

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Teresa L. Clark

Entitled
A SUBCONTINENTAL RECONSTRUCTION OF INVASION PATTERNS AND PROCESSES FOR THE PAST TWO CENTURIES

For the degree of Master of Science

Is approved by the final examining committee:

Songlin Fei
Chair

Jeffrey H. Holland

Bryan C. Pijanowski

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy of Integrity in Research" and the use of copyright material.

Approved by Major Professor(s): Songlin Fei

Approved by: Robert K. Swihart Linda S. Lee 11/30/2015
Head of the Departmental Graduate Program Date

A SUBCONTINENTAL RECONSTRUCTION OF INVASION PATTERNS AND PROCESSES FOR THE PAST TWO CENTURIES

A Thesis

Submitted to the Faculty

of

Purdue University

by

Teresa L. Clark

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science

December 2015

Purdue University

West Lafayette, Indiana

To my freakishly amazing daughters Reeves and Laurel,

I dedicate this work to you. Your resilience in the face of the upheaval of moving during your teen years shows a strength of spirit you should be proud to carry into adulthood. Through each of your own challenges you have gained knowledge about yourselves and accomplished more than you ever dreamed you could. I am grateful that in spite of my absences and preoccupation you both have matured into strong, independent women with unique talents and abilities. Although this work is distinctly my own, each page, each word belongs just as much to you.

ACKNOWLEDGEMENTS

My husband and children provided the support, love, food, and coffee that made this thesis emotionally and physically possible. Data, photos, time, and occasionally dinner, was provided by dedicated herbaria curators. Carol Ann McCormick at University of North Carolina gave me hope that this project could become a reality by providing me with specimen records on a weekly basis as their collection became digitized. Oliver Broenniman provided the stimulus and R script for the minimum dispersal routes reconstruction. Lab mates Johanna Desprez, Basil Iannone, Gabriela Nuñez-Mir, and Keeli Curtis helped with editorial and analytical advice. Graduate committee members Jeffery Holland and Bryan Pijanowski provided advice, direction, and inspiration. Funding was provided by a USDA NIFA Grant # 11207356, and from the NSF Macrosystems Biology Project # 1241932. Finally, Songlin Fei believed in me and gave me an opportunity to take on an immense project that only became a reality with his overabundance of patience and guidance. Thank you.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Clark, Teresa L.. M.S., Purdue University, December 2015. A Subcontinental Reconstruction of Invasion Patterns and Processes for the Past Two Centuries. Major Professor: Songlin Fei.

Understanding large-scale invasion patterns and processes is essential for effective and proactive management of exotic species that have caused significant ecological and economic damages. While many studies have focused on specific habitats and species, my research is aimed at the broad-scale patterns of expansion during the course of the past two centuries. I use long-term herbarium specimens to recreate the spatiotemporal distributions of 29 invasive plant species within the eastern United States. The emerging spatial patterns reveal species' expansion through a continuum of radial expansion, long distance jump dispersals, and range infilling. This study identifies a set of metrics that is able to quantitatively measure the changing morphology of the spatiotemporal patterns that emerge at macroscales. Utilizing graph and network theory I additionally reconstructed the most plausible processes of the invasion expansion. Results show that invasion is closely facilitated by human activity and there are considerable commonalities of distribution and dispersal patterns among invasive plant species.

CHAPTER 1.

"No matter where its seed fell, it made a tree which struggled to reach the sky.
It grew in boarded-up lots and out of neglected rubbish heaps, and
it was the only tree that grew out of cement….
That was the kind of tree it was."

-- Betty Smith*, A Tree Grows in Brooklyn (1943)*

1.1 Introduction

The single overarching principle that contributes to the expansion of invasive species is the human admiration for beauty, strength and resilience (Baker 1974; van Kleunen et al. 2015). Just as Tree-of-Heaven in the children's storybook could grow out of cement, many of the twentieth centuries' worst plant invaders exhibited the ability to grow where native species could not. The known ecological and economic effects of these invasions has produced much interest into predicting invasion patterns. Despite the large quantity of research over the past 50 years since Elton's seminal book (Elton 1958), we are still lacking the ability to form broad generalities of invasion at larger spatial scales. At these scales, humans have the ability, through our own movements, to influence the general invasion patterns that species exhibit. Recreating the historic

distribution of many different species can help us discern the implications of the cascade of individual smaller scale inputs. Historical collections housed in herbaria can be used as primary occurrence records to reconstruct the long term spread of invasive plants and study the emerging patterns. This study aims to utilize this wealth of untapped information to investigate how human mediation influences species invasion patterns and any departure within those patterns from accepted theories of natural species distribution and dispersal.

1.2 Background

The Millennium Ecosystem Assessment of 2005 identified invasive alien species in the top five direct causes of global scale biodiversity and ecosystem services decline (Millennium Assessment Board 2005). Spread of non-native species has been shown to not only detrimentally affect current ecosystem functions, but may have compounding effects on long-term species biodiversity at many scales (Charles and Dukes 2007; Pyšek et al. 2012). Local effects have encompassed species extinctions, changes in soil chemistry, and economic loss (Pimentel et al. 2005). Even though at regional scales biodiversity possibly increases in invaded areas, there is the likelihood for large-scale biotic homogenization (Ellis 2013; Powell et al. 2013).

Recreating the historic growth of the distributions of different species can provide us the tools to answer the more complex questions concerning the drivers of invasion. Gaining further understanding of the drivers of change is the critical link between the

invasive species inputs and the transformation of ecosystems (Bauer 2012). To fully grasp the complexities of the processes we need to understand how changes within the species, environment, and their interactions have varied through time (Strayer et al. 2006). Additionally, species distributions and invasions are dependent on scale (Levin 1992; Milbau et al. 2009) and studies of a single species or time period can impede our ability to gain deeper knowledge of cross-species and cross-scale interactions (Peters et al. 2007; Soranno et al. 2014). Therefore, studies with larger species groups that incorporate temporal changes have the possibility of producing more complete assessments of significant ecological impacts (Pyšek et al. 2012). This study addresses this need by using 29 plant species that were collected over two centuries throughout the entire eastern US.

1.3 Data Used in the Study

1.3.1 Species Choice

Invasive plant species are not confined to certain life forms or botanical families, and while they have common spread capabilities their particular traits are highly varied. To adequately cover the majority of invasion dynamics I began by choosing terrestrial plant species that were monitored by the USDA Forest Service in the eastern US. The full collection needed to contain representatives from each of the major life forms (i.e., trees, shrubs, grasses, forbs, and vines) and have large enough distributions to provide sufficient data. Additionally, to effectively cover the longest time period and possible

environmental interactions, I chose species that had an introduction history from the eighteenth to early twentieth century and had functional traits and life histories that differ from other invasive species (Appendix A). I chose these attributes because they are shown to likely affect the process of invasion (Mihulka and Pyšek 2001).

1.3.2 Herbaria

Long-term floristic data collected in herbaria can be used to reconstruct historic distributions by supplying the locations of known species occurrences, thereby providing the spatiotemporal progression of the invasion. Having this distributional data consolidated for use can enable research to proceed more rapidly and positively affect the management of biodiversity and ecosystem services (Meyer et al. 2015). Although this type of data contain collection biases, herbaria records have been shown to be a viable proxy for species range and expansion data and have been used extensively to study invasive spread (Aikio et al. 2010a; Barney 2006; Delisle et al. 2003; Fuentes et al. 2008; Pysek and Prach 1995).

In order to understand the nature and uses of herbaria data one must fully grasp how the collections were created and maintained. Herbarium records have been used extensively throughout history for different types of botanical studies. In 2003, Dr. Vicki Funk, curator from the US National Herbarium, published an article entitled "100 Uses for a Herbarium (well at least 72)." Number 38 on her compiled list was "track introduction and spread of invasive species (ecology)" (Funk 2003). Even prior to Dr. Funk publicizing this use, invasion ecologists and researchers had been using these

storehouses of historical records to perform research. Lavoie (2013) reported there were at least 382 papers published from 1933 to February 2012, that used herbarium specimens as their primary data. Many peer-reviewed study topics fell under the broad umbrella of plant and disease specific studies, but further topics ranged within climate change, pollution, and plant invasions. The author reported that almost 5 million specimens have been used in studies documenting biogeographical patterns; my thesis work additionally compiled another 40,000 specimens.

In the earliest days of the USA, herbarium such as the country's oldest, Salem College, which began in 1772 (Radford et al. 1974), held dried and mounted plant specimens on open shelves in natural air. As time progressed and specimens deteriorated, newer techniques required mounting to be on alpha cellulose or cotton rag paper, which is acid free and pH neutral. These herbarium "sheets" were then either frozen for a time or fumigated for insects, placed in protective folders, and enclosed in dust free cabinets. Many herbaria at this present day still house their collections in this manner, although most have adopted measures of climate control and compactor cabinets. These two sided shelving systems are movable on tracks to allow compression of the cabinets which enables the facility to house more specimens in a smaller space.

The cataloging of specimens has advanced as well, from no organization to worldwide digital databases. Since the time of written herbarium records, the arrangement and classification of the folders which contain the specimen sheets has not changed considerably, and most retain some form of the *Dalla Torre and Harms*

numerical arrangement of the Englerian System (Radford et al. 1974). As collections grew, the specimen sheets were additionally referenced by accession numbers recorded chronologically which allowed tracking of individual deposits.

Due to repeated use by students and researchers, curators needed to address the continued deterioration of the specimens from handling (Flannery 2013). Creating accessible records for those who do not need to touch the plant has become a priority for both the integrity of the plant material, and the furtherance of scientific inquiry (Suarez and Tsutsui 2004). Digitizing the collections extend both the quality and the usability of the specimens. Even within my tenure spent collecting specimen records, new digitized herbarium collections have become available, as multiple smaller herbaria join their collections together to form regional networks (Table 1.1).

Table 1.1 List of Consolidated herbaria networks that aid specimen digitation

Herbaria Network	**Abbreviation**
Southwestern Environmental Information Network	SEINet
Intermountain Regional Herbarium Network	IRHN
Southeast Regional Network of Expertise and Collections	SERNEC
Northern Great Plains Regional Herbaria Network	NGPherbaria
North American Network of Small Herbaria	NANSH
Consortium of Midwest Herbaria	n/a
California Consortia	CCH
Pacific Northwest Herbaria	PWH
Northeastern Herbaria	CNH
Texas Oklahoma Regional	TORCH
United States National	US
New York Botanical Garden	NYBG
Missouri Botanical Garden	MBG

1.3.3 Collecting Specimens

By accessing the online consortia websites and larger digital databases I easily collected many specimen records. To discover other herbaria, I accessed the New York Botanical Garden's (NYBG) resource *Index Herbariorum: A Global Directory of Public Herbaria and Associated Staff,* which provides herbarium name, address, location, collection size, and the curator's name of any registered public herbaria worldwide (Thiers [continuously updated]).

Digitized herbaria data available on the internet come in many forms which I needed to access and record differently. The simplest records I found to obtain were from databases which allowed exportation of the chosen specific fields of specimen records as either a text file or a spreadsheet. Not all the available herbaria that digitized their collections were able to procure the funding and personnel necessary for placing records on the internet. Many of the smaller herbaria have been able to have portions of their holdings archived within their in-house databases. These locally accessed data presented a challenge in that they have been recorded in non-standard formats and were not readily joined into the larger more standardized compendia.

Often for non-digitized collections, I found that a faculty curator or a student from facilities that housed smaller collections, was willing to photograph the species' specimens and send them to me electronically. Unfortunately, as collection size increased, many of the universities' faculty and staff could not spare the large amount of time to accomplish this task. To obtain resources from geographical areas not

represented by any digitized collections, I made site visits to eleven herbaria to photograph needed specimens (See list Table 1.2). Additionally, I was able to contract a student at Cornell University herbarium to photograph their specimens. The state of Kentucky has an absence of large or digitized herbaria, yet this deficiency was alleviated with previously collected samples by Liang (2010). Region wide specimen records for *Alliaria petiolata* (garlic mustard), including the earliest specimen collected, I obtained from Victoria Nuzzo from her research (Nuzzo 1993).

Table 1.2 Eleven University herbaria received site visit to photograph specimens

Herbarium	**ID**
University of Georgia	GA
University of Tennessee-Knoxville	UTK
Duke University	DUKE
Notre Dame University	NDG
Purdue University	PUL
Virginia Polytechnic Institute	VPI
West Virginia University	WVU
Lynchburg college	LYN
George Mason University	GMU
Virginia Military Institute	VMI
West Virginia Wesleyan College	VWC

The visits to herbaria provided their own challenges in the locating and handling of specimens. Through time the nomenclature of certain species has changed, and as such, different herbaria listed the same plant under different families. For example, I found the species *Paulownia tomentosa* filed within the families *Paulowniaceae, Scrophulariaceae* and *Bignoniaceae.* The organization of specimens in each facility was governed by the schema put together by the first curator and was usually a variation of the *Engler* system mentioned above. Specimen usage required time commitment and

care in handling. When removing samples from the cabinets I needed to take the entire folder to a separate counter to handle each specimen individually taking care to not invert, fold, drop, or slide it across another specimen. Depending on the size of the collection and time constraints, I photographed either the entire sheet or just the label area. For further verification or information retrieval, I archived all photos with the transcription data.

1.3.4 Associated Bias in Herbaria

Using historical collections requires an understanding of the time periods from which the specimens were acquired. Whether using museum manuscripts, natural collections, or archived photographs, all historical data collections have their own unique predilection for biases and sampling errors (Graham et al. 2004). History cannot be changed or overwritten, but must be accounted for within any scientific inquiry. Herbaria are set-up and maintained by humans with living histories, in consequence, their organization and function are subject to the whims of finances and world events. Collecting specimens non-randomly in relation to location, time, and taxa produces spatial, temporal, and species biases that must be addressed during study design (Pyke and Ehrlich 2010).

For example, during times of war or economic downturn the collections suffer both from lack of maintenance and decreased collecting effort and subsequent future rebounds can later cause large fluctuations in types and quantities of specimens added. Additionally, cross-taxa bias emerge because species which exhibit showy flowers are

collected more highly than more non-descript plants. Plants which are difficult to obtain due to height or that are challenging to mount on paper may also become less desirable to collect.

The temporal biases exhibited in this collection are apparent when graphing the quantity of specimens added to the database per year. The count of specimen numbers shows increased collection intensity occurring approximately in late 1950's and 1960's, then again near the turn of the century (Fig. 1.1). Mapping the distributions of all species in the database enabled me to visualize and evaluate locations where the highest concentrations of invasive species were collected, compared to the herbaria locations (Fig. 1.2). One of the notable causes of this trend in both spatial and temporal bias is the field work which produced the *Manual of the Vascular Flora of the Carolinas,* where within the first collection season in 1956, over 52,000 specimens were collected (Radford et al. 1968; Radford 1984).

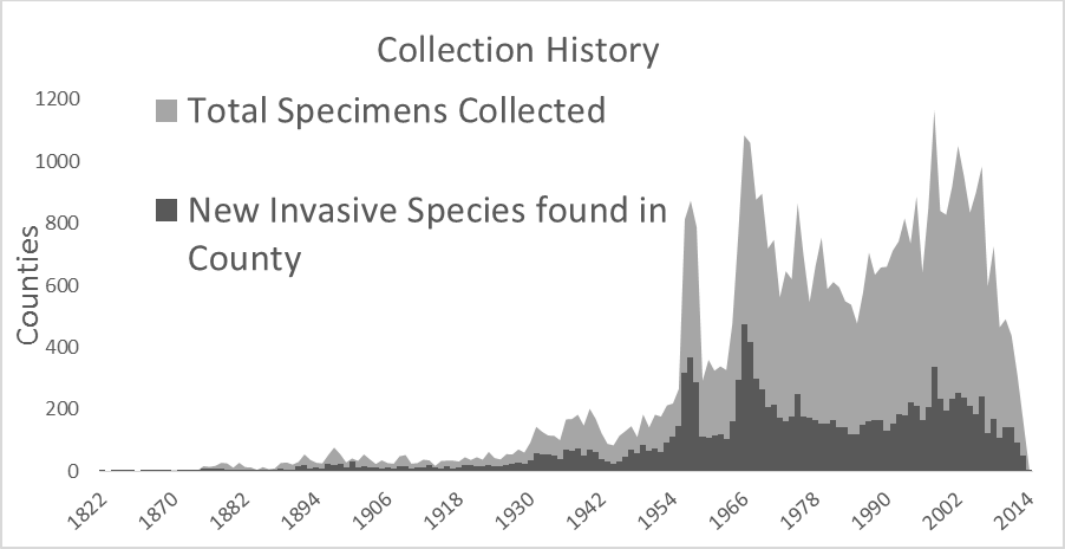


Figure 1.1 Quantity of herbaria specimens through time

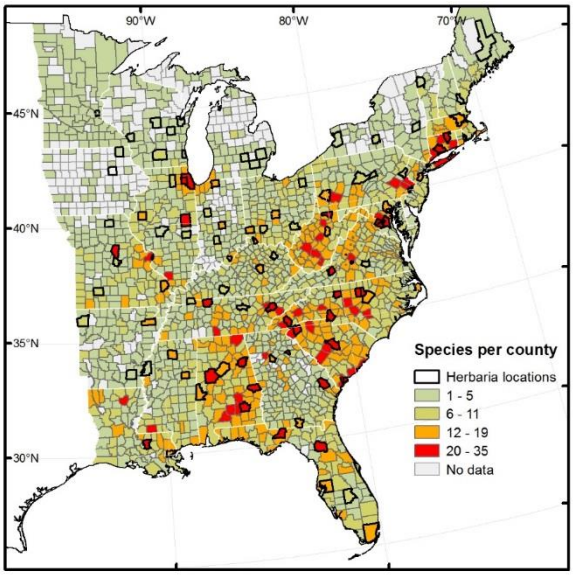


Figure 1.2 The total number of invasive species recorded based on herbarium records in each county

When using records from herbarium specimens to conduct research certain additional safeguards must be instituted for any conclusions to be broad reaching and valid. Previously, to alleviate some of these biases when reconstructing invasive species distributions, many researchers have included native species as a control within the studies (Crawford and Hoagland 2009; Delisle et al. 2003), or used overall collection intensity for all species as a gauge for accuracy (Chauvel et al. 2006), while still others extended null model approaches (Aikio et al. 2012). I chose to use the null model approach within Chapter 3 and only compare the species to its own sample distribution and not attempt direct comparisons across species. The utilized model also provided a measure of temporal stochasticity within its design. Within Chapter 2, the species invasion curve as it stands without any adjustments for collection intensity, likely over represents the rate of spread during the 1950-1970 era. *Lonicera japonica* was probably

existing in many locations the decade prior to collection during the period of World War II. These variabilities can be challenging to account for when designing a scientific study, but the information contained within the collections themselves are invaluable.

1.4 Data Processing

1.4.1 Database

The ability to compile many different data sources into one final database requires forethought concerning the further uses, naming conventions, and capacity to extract information without cumbersome guidelines. My project only required data to be recorded at the county level, and in the year of collection. I designed an initial common template for all obtained records that consisted of: (1) Institution identifier (from *Index Herbariorum*), (2) Species Code, (3) Family, (4) Genus, (5) Species, (6) State, (7) County, (8) Year (named Date due to ArcGIS requirements), (9) Common Name, (10) Acquired. In order to limit the final database to this information, I needed to compress or discard some information from many of the digitally obtained records. I could alternately have used the *Darwin Core* (Darwin Core 2004) schema which is used in many compiled biodiversity databases (Robertson et al. 2014), although it has a large amount of extraneous data which was unnecessary for my objective. Therefore, I took care to save and archive all original data without alterations. The records were initially collected in MS Excel, and later aggregated and transferred to MS Access.

Utilizing accepted techniques for herbaria specimens (Ahern et al. 2010; Crawford and Hoagland 2009; Delisle et al. 2003), I compiled the records in a database while maintaining the ten column headings. In addition, I obtained herbarium demographic data from *Index Herbariorum* to enable cross referencing the specimens to the specific institution from which they were obtained. By using differing queries I was able to filter the records to ascertain the first specimen for each species within each county, tabulate the accumulation of species and county specimens, and determine the total quantity of specimens received from each herbarium. The species' cumulative number of counties has been previously been shown to be an estimate of expansion rates (Pysek and Prach 1993).

1.4.2 Distribution reconstruction

I used the collected herbarium species presence records to reconstruct the spatiotemporal distributions for all 29 species. Records originated from approximately 200 different herbaria located widely across my study region (Fig. 1.3 & Appendix B). These specimens spanned a 192-year period (from 1822 to 2014) allowing me to analyze the county-level establishment and expansion over time.

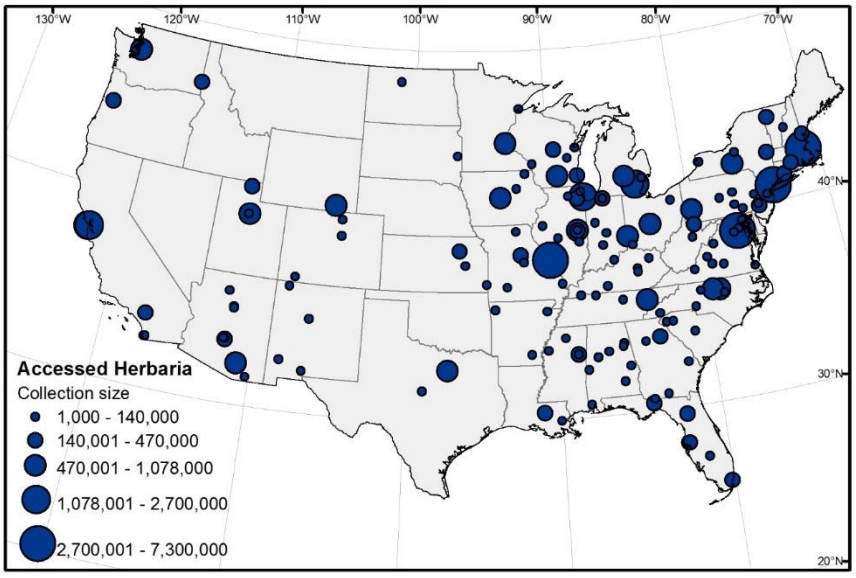


Figure 1.3 Location of US herbaria that were used in this study, and their collection size

After I graphed the cumulative number of specimen records (i.e., invasion curve) for each species, the species appeared to fall into temporal expansion groups (Fig. 1.4). Unlike the other three, the first grouping (Fig. 1.4a) mostly experienced long-term, slow, linear growth. Group 2 (Fig. 1.4b), spread to the largest number of counties and all experienced a period of exponential growth during the period 1950 – 1970. This same dynamic was repeated within the third group (Fig. 1.4c) although it was at a smaller degree and occurred for a shorter time period. The final group of species (Fig. 1.4d) had very slow growth from establishment until late within the twentieth century, then expanded rapidly.

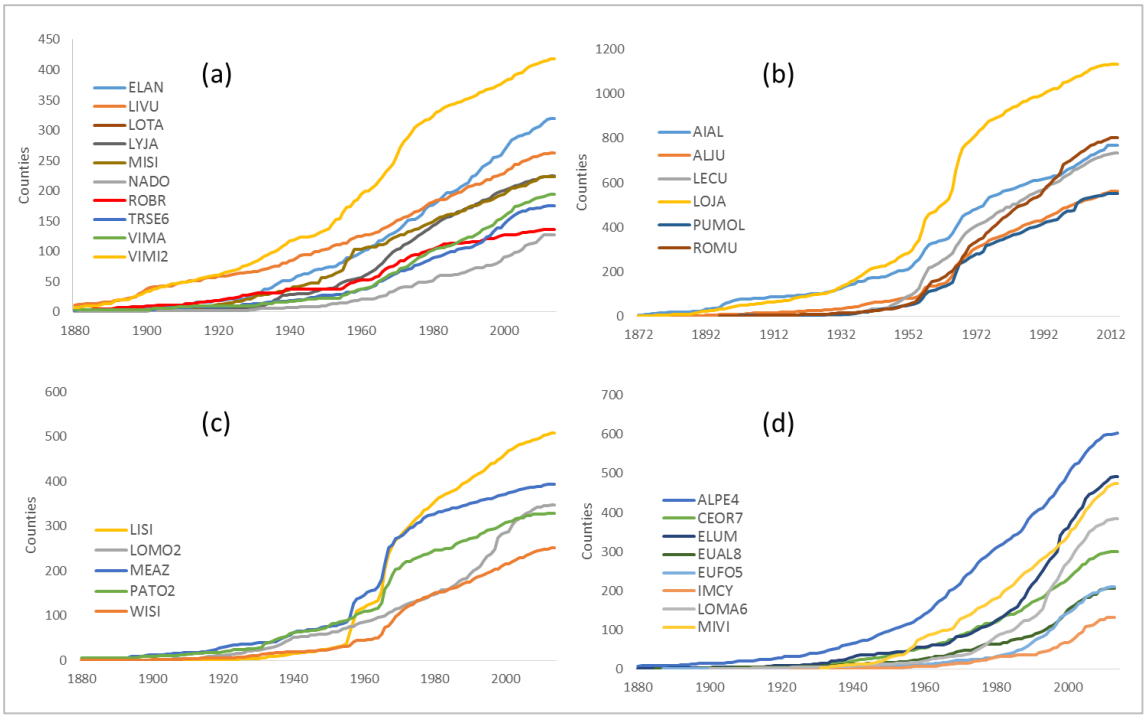


Figure 1.4 Cumulative number of counties over time for all 29 species

The earliest-occurrence county-specimen records were joined to the US census county polygon layer in ArcGIS (Esri, Redlands, CA). To facilitate point-pattern analysis, the species occurrence polygon layer was converted from county boundaries to the centroids of the occupied counties. I calculated all further analyses and measurements using the Albers Equal Area Conic Projection. Initial species network analysis (Chapter 3) utilized latitude and longitude coordinates and was completed with the World Geodetic System 1984 (WGS84) coordinate system.

I mapped the species temporally, which enabled viewing of the expansion visually, but it was difficult to ascertain any cross-species commonalties (Appendix C). Spatially, there are only two species that have current distributions uniformly covering most of

the eastern USA, *Rosa multiflora,* and *Lonicera japonica.* The other twenty seven species were either confined to one of three geographic areas: (1) the Southeastern states, (2) Northeastern and Midwestern states, or (3) a few which grew out of the Central states.

After static temporal mapping of the species distributions (Appendix C), I created Audio Video Interleave (AVI) movies with ArcGIS software to complete visual inspection. Using this time-series imagery I perceived three distinct patterns emerging throughout the species residence. Some species distributions began with counties speckled throughout their ultimate range, then continued to in-fill. Other species distributions grew out of one area with expansion appearing to occur in a wave-like front. Finally, there was a group whose distributions grew out of multiple different locations and eventually, through time, these separate divisions joined together. These classes of patterns provided the impetus for the developed toolset within Chapter 2, which can be used later to investigate species trait – pattern associations.

1.5 Research Questions

To approach the goal of identifying and evaluating emergent large-scale expansion patterns and their association with human mediation, I addressed two different facets of invasion patterns. Spatial patterns contain not only the objects arranged into a perceptual structure, but also the space between those objects. To confront this dichotomy I evaluated both the outwardly produced pattern morphology and its

temporal progression, and also the pathways of probable expansion that the species used to produce those patterns.

Measuring macroscale spatial patterns has historically been done through qualitative description of biogeographical movements. Guided by the hypothesis that invasive plant species distributions expand through a continuum of radial growth, longer distance jump dispersal, and range infilling, in Chapter 2 I identify a set of metrics that can be used to measure these patterns temporally. This set incorporates both distance-based spatial statistics and density-based landscape metrics by using point pattern analysis and kernel density estimation. The toolset was constructed within the guidelines of maintaining minimal data requirements and straightforward calculation and interpretation. The results show that a small set of well-chosen metrics can be used to measure the essential characteristics which define natural versus human-mediated species dispersal at macroscales.

Chapter 3 investigates the hypothesis that invasive plant expansion pathways through range interior networks, are associated with human mediation. A new model, which uses graph theory, was applied to the species occurrence records to reconstruct the most probable pathways of their expansion. By constructing a null model of permuted random occurrences I could study the variability between the proposed actual dispersal routes to that of random dispersion. Further comparisons between the species routes and human transportation corridors helped illuminate association between the two networks. Results show repeated use of specific expansion routes and locations that approximate a hub-and-spoke class of network structure. Those networks

additionally show strong association with the established human highway network. In its entirety, this work can help convey a deeper understanding of the differences and commonalities within the spatial patterns that emerge from exotic plant invasions during the past two centuries.

CHAPTER 2.

2.1 Introduction

Patterns of ecological processes across large geographic areas emerge from complex, cross-scale interactions among many biotic and abiotic factors. This complexity inevitably contributes to macroscale patterns of species distribution and dispersal. Understanding the factors that contribute to these patterns has long interested various sub-disciplines of ecology such as paleoecology and biogeography (Dobzhansky 1950; Matthew 1914; Pianka 1966), and is more recently crucial for understanding macroscale patterns of a leading component of global change—biological invasions (Mooney and Cleland 2001; Ricciardi 2007; Strayer et al. 2006). As a beginning step of understanding what drives macroscale dispersal patterns, one must first be able to effectively quantify various characteristics of species expansion patterns both over time and across large geographic areas.

Species distributional patterns change via a series of diffusion and jump dispersal events (Shigesada and Kawasaki 1997; Wilkinson 2001), leading to the continued evolution of spatial patterns over time (See Fig. 2.1 for examples). Being able to quantify the variability in these changing spatial patterns, along with the directional movement and rates of spread, can help to reveal differences and similarities in

patterns at different invasion stages of a given species and among different species. The quantification of these patterns, when paired with further understanding of species traits and environmental conditions, can provide increased understanding into the factors contributing to species expansion (Rouget and Richardson 2003). Despite the many indices developed to quantify spatial patterns, there is likely no single metric able to completely encapsulate the complexities of large-scale dispersal patterns (Pysek and Hulme 2005). The quantification of spatiotemporal changes in species distributions across large geographic areas will likely require multiple metrics. Identifying these metrics, however, is challenging due to the lack of spatiotemporal data pertaining to species distributions.

The main objective of this investigation is to identify a set of metrics applicable to macroscale pattern recognition using methodologies already tested at smaller scales. The utility of these indices will be evaluated using spatiotemporal data for two invasive plant species, *Alliaria petiolata* (M. Bieb) Cavara & Grande (garlic mustard) and *Lonicera japonica* Thunb. (Japanese honeysuckle), that exhibit distinctly different macroscale patterns of spread. The outcomes of this investigation will provide a set of metrics that can be used at macroscales to quantify spatiotemporal patterns of species dispersal, which can aid better understanding of those patterns, and perhaps of other ecological processes. These findings will also provide a greater understanding of macroscale patterns of spread for invasive species, which is of practical utility given the contribution of these species to overall global change (Fei et al. 2014; Simberloff 2011; Vitousek et al. 1997).

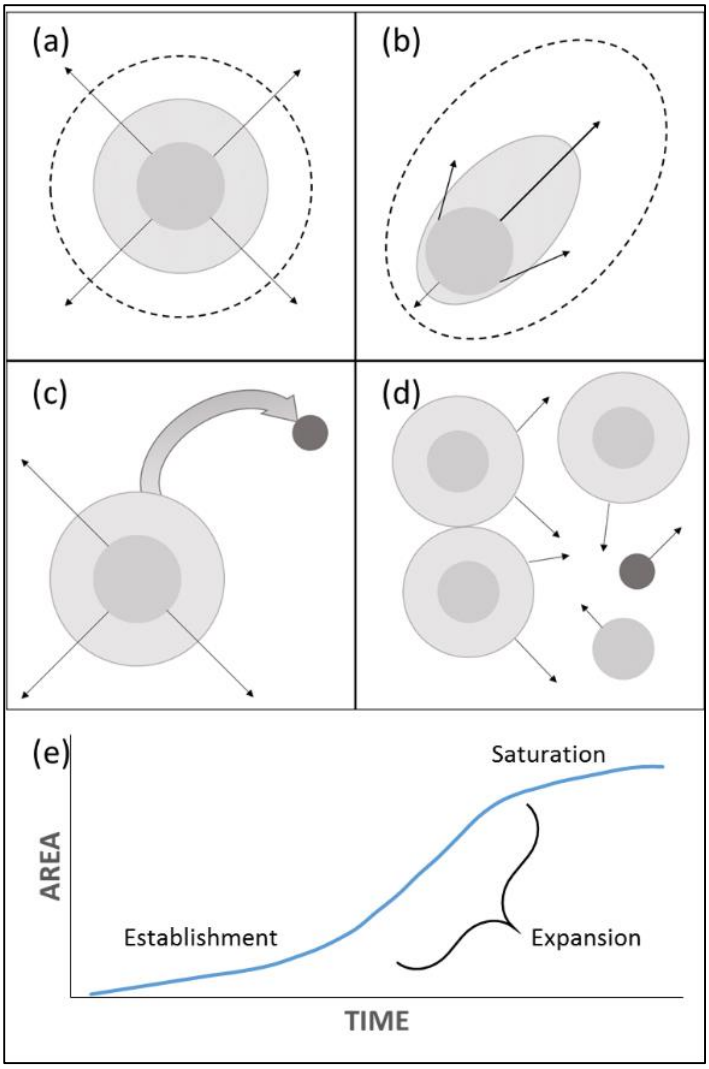


Figure 2.1 Typified patterns of species expansion, including (a) radial diffusion from initial point of introduction, (b) directional radial diffusion due to barriers or environmental constraints, (c) a combination of radial diffusion and jump dispersal events, (d) initial dispersal to multiple satellite colonies from which diffusion and range expansion occur, and (e) establishment, expansion, saturation phases of invasion as determined by changes in rates of spread

2.2 Materials and Methods

2.2.1 Data Selection and Acquisition

To estimate patterns of spread for both species, I determined the timing of county-level presence of each invader for the eastern United States of America (USA) by using herbarium records from the past two centuries. Each species exhibited variable rates of spread over the courses of their invasion histories, allowing me to identify different phases of their invasions (see Fig. 2.1e for definition of invasion phase). Specimens for *A. petiolata* and *L. japonica* were collected from approximately 120 differing herbaria and other published sources (Appendix B). Following protocols by (Delisle et al. 2003), specimens lacking locations, dates, or which were obviously cultivated were discarded. Herbarium specimen records have been shown to be a viable proxy for species range and expansion data (Aikio et al. 2010a; Pysek and Prach 1995), and have also been used successfully to reconstruct the timing and extent of invasion (Barney 2006; Fuentes et al. 2008; Kannan et al. 2013). While herbarium data is known to exhibit both spatial and temporal bias due to sampling effort being non-random (Loiselle et al. 2008), the purpose of our study was to described and differentiate spatial patterns and not to evaluate the actual species invasions, making these biases inconsequential.

Each of the species has a long invasion history in the USA, being introduced in the early 19th century (Munger 2001; Schierenbeck 2004). *Alliaria petiolata*, an upright obligate biennial forb, is believed to have been introduced for use as a garden herb and

medicinal purposes (Munger 2001). Due to its prolific seed production (Nuzzo 1999), and its primary mechanism of dispersal by epizoochory (Cavers et al. 1979), its recent spread is believed to be mostly accidental (US Forest Service 2005). In contrast, *L. japonica*, a perennial flowering vine, was introduced through the horticultural trade because of its fragrance and ease of growth (Lemke et al. 2011). The variability between these species in their growth forms, modes of introduction, and life history traits likely influences their invasion trajectories and current range sizes (Moravcová et al. 2015).

2.2.2 Metrics

Spatiotemporal dynamics of range expansion encompass a vast continuum of spatial patterns that lie within the general bounds of diffusion (i.e., gradual and regular spread, radiating in concentric circles) and jump dispersal (i.e., long distance spread to locations outside of a species current range) (Pielou 1979; Simberloff et al. 1997). To capture the entire gamut of potential spatial patterns in range expansion and infilling, it is essential to have metrics that measure both the density of, and the distances among, species observations. In addition, to determine if spatial patterns of expansion vary among invasion phases, we need a quantitative method for detecting abrupt changes in rates of spread, which likely signify points in time where these phases begin and end. Furthermore, to ensure the greatest utility and applicability of these metrics, our metric choices (Table 2.1) were weighed by ease of use, data requirements, independence from the other metrics, intuitiveness, and their ability to distinguish spatiotemporal patterns at the macroscale.

Table 2.1 Metrics used to quantify macroscale patterns of spread including their equations and specific use

Metric	Abbreviation	Equation	Purpose
Regression slope	SLR	$\beta_1 z_i + \beta_2(z_i - \varphi) + \cdots$	We used segmented linear regression to model the cumulative number of counties in which the species occurred in response to time. Used to detect changes is slope estimates and thus changes in invasion phase
Standard distance	STD	$SD = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{X})^2}{n} + \frac{\sum_{i=1}^{n}(y_i-\bar{Y})^2}{n}}$	Dispersion around the mean centre, a geographic standard deviation of points, i.e. county centroids, used to assess the degree of scatter for a distribution
Mean centre	MC	$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n}, \bar{Y} = \frac{\sum_{i=1}^{n} y_i}{n}$	Average x,y location of geographic center of a species distribution. Detects directional movement in consort with STD.
Number of patches	NP	Count	The number of kernel density patches. Aggregation is revealed by the decreasing number of patches indicating coalescence.
Coefficient of variance of Euclidean nearest neighbor	ENN_CV	$CV = \frac{SD}{MN}(100)$	Standard measure of relative variability of inter-patch distances. Spikes in value indicate jump dispersals.

To estimate the years at which the species' expansion rates changed, I performed a *segmented linear regression* (SLR) on the species expansion area (cumulative counties) versus time (year). This graph and its coinciding segmented regression thus depicted each species' invasion curve (*sensu* Pysek & Prach,1993). The final recorded number of counties invaded by the two species differed considerably (n=581 for *A. petiolata* and n=1075 for *L. japonica*). Therefore, I unit-base normalized the data prior to regression analysis to facilitate cross-species comparison using the following equation:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

To estimate the degree to which distributions were dispersed throughout a species' range and the directional migration of each species, I utilized two distance-based spatial statistic metrics: *standard distance* (STD) and *mean centre* (MC). I calculated these metrics from centroid locations of the counties from which herbarium specimens were collected. Two density-based metrics: co*efficient of variation of the Euclidean nearest neighbor* (ENN_CV) and the n*umber of patches* (NP), were then used respectively, to identify jump dispersal events and to estimate intra-range species dispersion. To calculate these metrics based on the density of counties in which each species occurred, I rasterized the county centroids with a fixed kernel density estimator (KDE) and 75 percent volume curves (PVC) were drawn (Geospatial Modeling Environment, Version 0.7.3.0) based on prior recommendations (Borger et al. 2006).

I also explored the utility of numerous other metrics (see Appendix D), but for multiple reasons found them to be less applicable to macroscales. I chose segmented regression over others techniques (Aikio et al. 2010b; Larkin 2012; Liebhold et al. 2013; Mihulka and Pyšek 2001) because of its ability to detect changes in rates of spread, and thus transitions between different invasion phases (e.g., Fig. 2.1e). I also measured point distances between the centroids of all invaded counties and evaluated the effectiveness of using the maximum, mean, and standard deviation of those distances as metrics of spatial dispersion. These metrics were not chosen, because relative to STD, they did not provide a significant improvement in pattern description and they required more data and computational time. Additionally, I investigated other metrics aimed at quantifying aspects of area and edge, shape complexity, and aggregation. I opted not to

utilize these landscape-level metrics because either they were incapable of detecting spatial variability in our presence/absence data at the much larger spatial extent of our investigation, or they were less intuitive.

Spatial metrics were calculated for decadal time intervals. Segmented regressions were conducted using the Segmented package in R (Muggeo 2008; R Core Team 2015). Kernel density estimation and all distance geographic metrics were calculated using the Albers equal area conic projection (USGS) and ArcGIS 10.2.2 (Esri, Redlands, CA, USA). Landscape metrics were calculated utilizing FRAGSTATS (v4.2.1.603) (McGarigal et al. 2012). The final set of five metrics were verified for independence (see Appendix E) and the four spatial metrics were normalized from 0 – 1 to facilitate cross-species comparisons (see equation 1 above).

2.3 Results

2.3.1 Thresholds and Rates of Spread

Assisted by segmented linear regression, I found three changes in rates of spread corresponding to two invasion phases for *A. petiolata* (Fig. 2.2a) and four changes in rates of spread corresponding to three possible invasion phases for *L. japonica* (Fig. 2.2b). The first two line segments for the invasion curves of both species likely signify establishment phases. This distinction is due to both the relatively lower slope values for these line segments and the paucity of herbarium specimens collected during these periods. The subsequent segments, where rates of expansion increase, likely signify

expansion phases. *Alliaria petiolata* showed a tripling in its rate of spread after 1960, which it still maintains. Comparably, the non-scaled counts show *L. japonica* accumulated 25 counties per year during its expansion phase versus *A. petiolata's* 9 counties per year, but *L. japonica* only maintained that rate of spread for approximately 20 years. After that period *L. japonica's* rate of spread declined, suggesting it is approaching a saturation phase. Segmented linear regression revealed comparable values in the rate of spread for both species during their establishment phase, and the species' shift to expansion phase occurred in close time proximity to each other. Yet, the differences in normalized rates of spread during expansion phase (0.015 normalized units for *A. petiolata* and 0.025 normalized units for *L. japonica*) allowed us to investigate whether varying rates of spread would affect the utility of our spatial metrics.

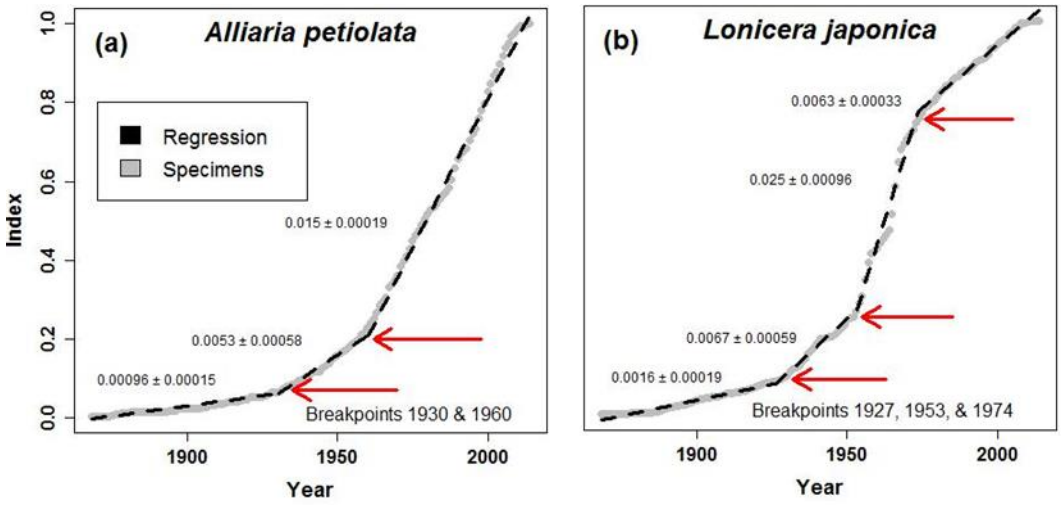


Figure 2.2 Results of segmented linear regressions with scaled slopes and se values revealing (a) an establishment and expansion phase for *A. petiolata* and (b) establishment and expansion phase for *L. japonica*, with this species potentially approaching a saturation phase

2.3.2 Diffusion and Jump Dispersal

The four distance- and density-based spatial metrics were able to quantify where along the continuum of diffusion and jump dispersal events occurred, and were able to differentiate these patterns between our investigated species. Temporal changes in STD values revealed the ranges of the two species exhibited contrasting levels of outward diffusion. STD is a measure of dispersion that helps to assess whether the increase in area occupied by a given species (in our case, counties) is due to outward expansion or infilling of the species' range. The metric estimates the radius of the circle that encompasses one standard deviation of our sample points, i.e., centroids of the invaded counties (68% assuming a normally distributed population). The continually increasing STD exhibited by *A. petiolata* (Fig. 2.3a & c) is the result of this species' distribution growing outward from a single foci located in the northeastern USA (e.g., Fig. 2.1b). On the other hand, *L. japonica*'s STD declines through the twentieth century (Fig. 2.3d) owing to the species distribution diffusing from multiple introduction points located throughout its geographic distribution (e.g., typified by Fig. 2.1d). When a species distribution infills more than it expands, such as with *L. japonica*, STD values decrease (Fig. 2.3b).

I found both species experienced initial increases, followed by decreases, in the degree of dispersion exhibited by their distribution. This pattern is revealed by the initial increase and then later merging of kernel density patches (NP) (Fig. 2.3 e & f). The NP metric allows investigation of the species' tendency to group into regional clumps,

which augments the STD's ability to distinguish overall long term changes in dispersion. Both *A. petiolata* and *L. japonica* show an increase in NP from their initial colonization until the maximum NP count occurs during the transition from the establishment to expansion phase (Fig. 2.3 c & d). Throughout their expansion phase, the NP values for both species continually decrease, showing patch coalescence into larger, less numerous patches. *Lonicera japonica's* NP values plummet, while those of *A. petiolata* maintain moderate values. The plummeting NP values of *L. japonica* are largely due to infilling of the species' range (Fig. 2.3f). The moderate yet stable NP values of *A. petiolata* are due to new patches occurring over time as this species' range expands westward, while older patches simultaneously merge together (Fig. 2.3e).

Utilizing the *coefficient of variation* from the measured Euclidean nearest neighbor distances (ENN_CV) calculated between the kernel density patches, I was able to detect instances of jump dispersal within the history of *A. petiolata,* but not in the history of *L. japonica*. The coefficient of variation is a ratio of dispersion that shows the extent of variability of the sample in relation to its mean (in our case, the measured distances at that time period), and is calculated by taking the standard deviation of a frequency distribution and dividing it by the mean of that same distribution (i.e., $s/\bar{x}$). Using the Euclidean nearest neighbor distances between all patches as our frequency distribution, abrupt spikes in the values of ENN_CV reveal the existence of an outlier patch, and thus a jump dispersal event. For example, before being normalized to facilitate comparisons between species, if the calculated ENN_CV of the measured distances is greater than 100% it indicates that the largest distance between patches is more than double the

overall mean inter-patch distance (McGarigal et al. 2012). The two spikes in value seen within *A. petiolata's* history (Fig. 2.3c) arise from measured ENN_CV values of 116% and 143%. These instances of high values signify the appearance of patches far away from all others, and thus an indication of jump dispersal, e.g., such as the outlying patches resulting from *A. petiolata's* westward expansion.

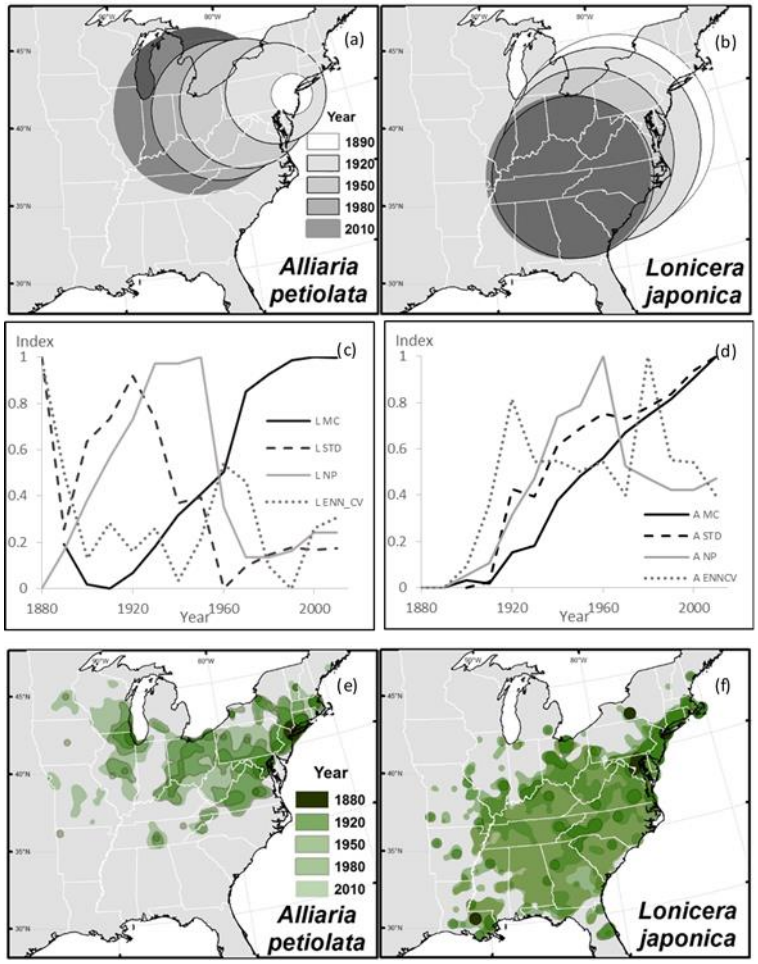


Figure 2.3 Differing spatial patterns between *A. petiolata* and *L. japonica*, as revealed by four different metrics. Shown are: mapped temporal changes in standard distance (STD) and mean centre (MC) for (a) *A. petiolata* and (b) *L. japonica*; temporal changes in all four spatial metrics (STD, MC, NP, and ENN_CV) for (c) *A. petiolata* and (d) *L. japonica*; and mapped temporal changes in the kernel density patches used to calculate NP and ENN_CV for (e) *A. petiolata* and (f) *L. japonica*.

2.3.3 Directional Expansion

The mean centre (MC) value for each species revealed the directional trajectory of its expanding distribution. The MC of *A. petiolata* shows a continual linear increase indicating persistent directional expansion, while that of *L. japonica* shows a rapid decline in directional movement after 1960 (Figs. 2.3a, b, c, &d). MC of the distribution is calculated from the average of the (x, y) geographical coordinates of all the input points (in our case county centroids) at each time period. While the continual migration of MC away from a species' introduction point indicates directional movement, a stationary MC requires further investigation. A stationary MC can either be resultant from range infilling, as seen in *L. japonica's* later years (Fig. 2.3e and typified in Fig. 2.1d), or indicative of equal expansion in all directions radiating from a species' initial introduction point. During *L. japonica's* recent history, and most recent invasion phase, all four spatial metrics (STD, MC, NP and ENN_CV) level off, while the rate of spread decreased (Fig. 2.2b), suggesting the distribution of this species to be stabilizing. Alternately, if a species is expanding equally in all directions, the MC would likewise not move, although the STD would increase.

2.4 Discussion

Detecting and measuring spatial patterns is a first step in achieving greater understanding of macroscale dispersal patterns. I was able to measure the essential characteristics of large-scale patterns of species distributions, including diffusion, jump

dispersal, and changes in rates of spread (Fig. 2.1) by using a group of five differing techniques and metrics (SLR, STD, ENN_CV, NP, and MC). When used in consort, these techniques and metrics proved capable of detecting both subtle and conspicuous differences between the spatiotemporal patterns of the two invasive plant species. Together, distance based (STD & MC), and density based (NP, ENN_CV) indices enabled us to quantify complex macroscale patterns, while the additional use of a segmented linear regression (SLR) enabled the division of species invasion histories into distinct invasion phases. This division was necessary to ensure that cross-species comparisons occurred during the same invasion phase.

Although simple comparing to other macroscale toolsets (Rangel et al. 2006), I feel this attribute underwrites the utility of the proposed metrics. Both simple and broad measures are often necessary in early exploratory analysis in many disciplines, and macroscale studies are likely no exception. Having a base-line method to apply to large spatial scales can allow macroscale research to proceed more rapidly, versus attempting to scale up results from smaller scale studies (Miller et al. 2004). The ability to quickly distill large datasets into a few common elements is becoming more critical as macrosystem databases are becoming constructed and available for use in research (Heffernan et al. 2014).

While sets of metrics like these used to analyze spatial patterns provide an introductory step in gaining a deeper understanding of the underlying ecology of large-scale spatial patterns (Wu et al. 2011), they also may provide a straightforward, standardized method. Having a standardized method for investigating overall patterns

can aid cross study comparison and further expose broad generalities (Cadotte et al. 2006). This recognition of large-scale spatial patterns can in turn guide further investigations into the causation of these patterns and a possibly provide greater understanding of interrelated and underlying processes.

Specifically, it may prove useful to apply these metrics to evaluate the effects of life history and functional traits on macroscale patterns of invasion. Some traits that aid species in establishment in novel environments may later hinder their expansion in later phases of invasion (van Kleunen et al. 2015). Therefore, the evaluation of the effects of species traits on spatiotemporal patterns needs to be invasion-phase specific. By utilizing a SLR to identify distinct invasion phases, we can more readily tease apart those traits that enable a species to be more invasive during each phase. Additionally, cross-species comparisons made during the same invasion phase may provide increased understanding of life history dynamics and may be of greater practical utility for prioritization of management actions (Van Kleunen et al. 2010; Williamson et al. 2005). Our investigation showed that *L. japonica* reached a far higher maximum rate of spread during its expansion phase than did *A. petiolata,* which is echoed by the larger vacillations in *L. japonica's* other metrics. During *L. japonica's* expansion phase the number of discrete patches (NP) plummeted to the same level seen in the late 1800's resultant from almost all of its expansion stemming from range infilling. Since that time this species has become ubiquitous in southern forests and is the most widespread invasive species in the southeastern USA (Oswalt and Oswalt 2011). The current slight rise in NP values (Fig. 2.3d) may be due to the emergence of new, western patches (Fig.

2.3f) signaling that the previous expectations of range growth (Schierenbeck 2004) are just beginning to be realized.

Although I used these metrics with long-term data, they may also be useful for differentiating macroscale patterns among more-recent invasions. While this set of metrics can certainly aid in the investigation of biological invasions, part of their value rests in their ability to be applied to a wide range of investigations pertaining to spatial patterns. Possibilities may include changes in vegetation cover obtained from time series of aerial imagery, monitoring of wildlife migratory ranges, or even evaluating trends within interdisciplinary research collaborations (Hoekman et al. 2010; Martin et al. 2015; Roush et al. 2007).

For continued research into macrosystem patterns and processes, ecologists need to be able to quantify the patterns that emerge from the spatial heterogeneity in ecological processes that occurs across large geographic areas (Heffernan et al. 2014). The utility of these techniques and metrics presented here for meeting this need lies within their minimal data requirements and their straightforward calculation and interpretation. Tools, such as these, can enhance our ability to detect and quantify biogeographical patterns, which is an important step towards further understanding the causes and consequences of these patterns.

CHAPTER 3.

3.1 Introduction

Humans value connections. Our brains are organized into neural networks (Fox et al. 2005) and our movement patterns follow predictable frequent routes (Gonzalez et al. 2008). Historically there has always been the well-worn pathways that marked the most traveled routes. These routes are not always the shortest distance between two points, but generally serve a purpose to direct the traveler to water, shelter, safety, or pleasure. As humans moved their homes on these paths through time, flora and fauna were carried with them for sustenance, comfort, and nostalgia (Baker 1974; Baker 1986; Heywood 1989). More recently transported species have often assumed the purpose of beautification or restoration of landscapes (Reichard and White 2001). When some of those species become invasive, they spread across the broad divisions of both natural and human mediated dispersal (Aikio et al. 2012; Ricciardi 2007). The intense spread of non-native species has been expedited by recent human commerce and transportation networks (Banks et al. 2014; Pysek and Hulme 2005). Whether accidental or purposeful, human mediated dispersal and transportation of species accelerates their spread and may lead to increased homogenization of species assemblages (Capinha et al. 2015).

To shed light on the different mechanisms involved within invasive species spread researchers have long been interested in recreating species historical expansion patterns (Pysek 1991). Previous recreations of invasive species expansion patterns have often used individually or in combination herbaria specimens, flora records, and historical reports to document the occurrence locations (Aikio et al. 2010a). These studies have depicted the species distributions through expanding grid networks (Pysek and Prach 1995; Weber 1998), county occupancy (Barney et al. 2008; Nuzzo 1993; Suarez et al. 2001), distinct location points (Lyons and Scheibling 2009; Mihulka and Pyšek 2001), and genetic analysis (Burrell et al. 2015; Durka et al. 2005).

Recently a model constructed to recreate species historical distributions that uses graph theory, has been shown to be viable in projecting the possible least cost pathways of dispersion of an invasive species (Hordijk and Broennimann 2012). Graph theory and network analysis can trace its beginnings to 1736 and the Königsberg Bridge problem which required one to devise a path through the city that would cross all 7 bridges only once and finish at the same place it began (Biggs et al. 1976). Highly utilized within the fields of mathematics, computer science, and electrical engineering, it has not been applied in ecology until recently. Graph theory has been shown valuable in modelling habitat connectivity (Koh et al. 2013; Minor and Urban 2008) and smaller-scale invasive species simulations (Ferrari et al. 2014), yet this approach is novel in studying invasive species at large spatial scales. This type of adaptive model may prove to be of great use in discerning large-scale temporal trends that may not be apparent with the use of only

discrete time period recreations as seen in Chapter 1 and in the studies mentioned above.

Least cost pathways of invasive spread at coarser spatial scales have generally been theorized to coincide with human transportation networks (Christen and Matlack 2009). Therefore, global industrialization and development of strong intercontinental and sub-continental network structures provide not only a means for species arrival and establishment, but also a supporting channel for their expansion (Hulme et al. 2008; Lavoie et al. 2007; Lelong et al. 2007; Liebhold et al. 2013; Wilson et al. 2009). Improvements to road systems not only disturb the local species associations and encourage the roadway edges to become highly invaded, but additionally can function as a conduit for the invasion of landscape interiors (Gelbard and Belnap 2003).

To expand on these premises and to explore further the human mediated effects on the expansion network dynamics of invasive plant species, there were three driving questions that guided my research: (1) Do the distributions of the invasive plants in my study expand from the closest occupied area as would be expected from a reaction – diffusion model (Skellam 1951), or from long distance dispersal (i.e., human mediated) as portrayed by the stratified – diffusion model (Shigesada et al. 1995)? (2) Are the species' networks and connectivity between nodes formed from natural, random, or organized processes (Banks et al. 2014)? In other words, do most counties have few outgoing edges while others function as *invasion hubs* that are the major sources for the expanding invasion? (3) Are the expansion network routes that connect the county

centroids more associated with human constructed major US highways than would be expected from connectivity that evolved from random expansion?

The overarching goal of this study is investigate the hypothesis that invasive plant species' expansion pathways are associated with human mediation. By studying these larger, regional-level invasion patterns the results may prove more revealing of robust broad-scale relationships between human mediation and invasive spread (Lonsdale 1999; With 2002). This research will show how reconstructing species distributions with long-term data can be approached with integrative and adaptive methods that provide a novel perspective concerning invasive species expansion patterns.

3.2 Methods

3.2.1 Data

This study uses the previously collected and described herbaria data (See Chapter 1) with the first occurrence record in each county for each species. Due to the difficulty obtaining actual geographic coordinates for each specimen I used x-y coordinates for every centroid of the occupied counties as approximates to study species spatial movement at the regional level.

The US roads layer was obtained from the Federal Highway Administration's National Highways Planning Network (NHPN) roads database (Federal Highway Administration 20140501). This geospatial line layer regards all types of road segments equally without consideration of travel volume, road length, or full versus limited

access. As I aimed to investigate the association of species' expansion to the major transportation corridors, the major roads (e.g., Interstates, U.S. highways, and State highways) were segregated out of the initial line layer. To be able to identify the species' longer-distance dispersal association with the mechanisms of human travel, I weighted the limited-access interstates highest with a level of 3 due to their ability to facilitate long distance dispersal. The US highways that were long distance but provided free access provide more on and off traffic, I therefore weighted them as 2. The mesh of short-distance, free access, two lane State highways was allotted a weight of 1. This technique was additionally deemed necessary as species network edges follow direct lines between county centroids and may lay parallel to the connecting highly traveled, limited access highways without physical coincidence. Using this weighting system and the line density tool in ArcGIS, the road layer was converted to a raster (cell size 9384 m^2).

3.2.2 Theory

In the field of mathematics, a graph is a set of individual entities (i.e., nodes or vertices) connected to each other by some relationship (i.e., edge) (Easley and Kleinberg 2010). If the relationship flows from one node to another it is called a directed graph, as exemplified by a person throwing a ball to another. If the relationship is common and symmetrical to both nodes it is an undirected graph, as two people shaking hands. Spread of invasive species functions as a directed graph, because when the plant expands into a new county it emanates from the first; each county experiences the

relationship differently. Additionally, directed graphs can be described as trees or arborescence such as those used to construct phylogeny of species. Each possible edge of the graph contains an intrinsic value of the amount of weight or *cost* for the relationship to travel that pathway.

Utilizing a directed graph and a specified cost function for each connection (edge) between counties (nodes) we can construct a *minimum cost arborescence (MCA)* or also called a *minimum spanning tree (Bock 1971).* The solution to this problem produces a minimum cost for invasion of a species to travel to all time ordered nodes without any backtracking or circularity. In other words, when a county becomes invaded, that outgoing edge may not reverse flow to the county of the invader, nor create an enclosed circuit back to any previously occupied county.

Due to the uncertain nature of occurrence records (i.e., it is probable the location was occupied for some time before a specimen was collected), it is undesirable to accept the first MCA output as the only possible. If a small random value is subtracted from each of the known times it will introduce a level of uncertainty and provide that each county-time value is unique. By having all input times as unique values, the graphing problem reduces in complexity and can be more easily solved with a *greedy* algorithm (Hordijk and Broennimann 2012). By definition, a *greedy* algorithm is one that aims to obtain a global optimum by making the optimum choice at each independent step in the process. To begin construction a MCA requires a starting node that will not have any incoming edges. As alien species often have multiple introduction points, the

graph can alternately be constructed with more than one initial point and will produce separate unconnected networks for that species emanating from the different nodes.

The evaluation of a graph's connectivity is part of network theory, which examines the intervening relationships between the nodes. Network research has played an integral role in analysis of the internet, social networks, journal citations, and most recently molecular level protein interactions and gene expression (Hauff and Nürnberger 2005). Beyond uniform grid-like graphs, the two main types of disordered connectivity are considered to be random and small-world networks (Fig. 3.1). Each type and individual graph has a *degree* which is the number of edges meeting each vertex. Random graphs (Fig. 3.1a) tend to have degree distributions that approximate a Poisson distribution (Fig. 3.2), with most vertices having mid-level connectivity. Small-world networks are more highly connected with shorter length edges and may have clusters of more highly connected areas (Fig. 3.1b). Within small-world networks there exists a class of scale-free networks (Fig. 3.1c) that have a few, significant number of nodes with many edges and a long tail of nodes with few connections (i.e., as Fig. 3.2 power-law curve) (Barabasi and Albert 1999). The structure is considered scale-free because the connectivity is not altered if the network is viewed or evaluated at different scales (Amaral et al. 2000). A scale-free network follows a power-law distribution of node linkages that is long-tailed and appears as an asymptotic hyperbola which becomes linear on a log-log plot (Barabasi and Bonabeau 2003). A power-law distribution occurs when the frequency of an event (in my case a node having a specific

number of outgoing edges) is related to the size of the event and uses a constant exponent (λ) which defines the decay shape of the distribution (White et al. 2008).

$$P(k) \approx k^{-\lambda} \quad (1)$$

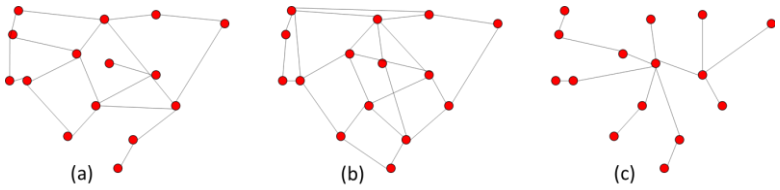


Figure 3.1 Classes of graph networks (a) Random networks have connections that are randomly placed (b) small-world have shorter length edges and areas of clustering and, (c) scale-free networks have node connections that have hubs with a high number of linked edges

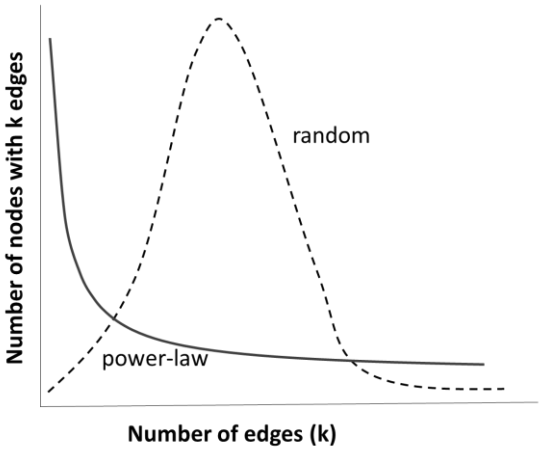


Figure 3.2 Frequency distribution of nodes with certain number of edges. For random networks the number of edges for each node follows a Poisson distribution; for a power-law distribution the number of nodes decreases constantly as edge number increases (i.e., there are few nodes which function as hubs with many connections)

3.2.3 Computing Script

The R script provided by Olivier Broenniman (Broennimann et al. 2014; Hordijk and Broennimann 2012) uses a MCA function and returns a minimum dispersal route (MDR) for the species when provided with years and locations of species occurrences. The source code contains options for the production of MDRs for minimum possible distance

routing and random network routings, beyond the expected network from the actual observed dates. To compute the minimum route, the dates of the specimen occurrences are disregarded, and the shortest directed network is formed that incorporates all the points. The sample network distribution is constructed by randomizing the occurrence years across all the locations instead of using the specimen's actual date. By holding the locations fixed and randomizing the assigned dates it reveals how the species could have reached all the locations by chance.

To add a stochastic component and uniqueness, the small random value to be subtracted is drawn from a truncated negative exponential distribution (95% quantile) with a user specified mean. Other user supplied parameters include the quantity of initial nodes and the number of times to iterate the function. If the function is repeated more than once the output will contain bootstrap values that show the confidence level of that segment (i.e., edge). For example, when I iterated the process 100 times, a route that occurred during all iterations would receive a value of 1, while a single occurrence of that route would return 0.01.

3.2.4 Process

The supplied R script was altered slightly to provide for more streamlined integration into ArcGIS through Model Builder ®, by removing unusable characters in the column headings and adding a species identifier to each output segment.

User specified parameters were chosen based on the aggregate histories of the collected plant species. For model standardization and consistency across the entire

database of species, I chose my initial nodes to be the first two occurrence records for each species. Choosing two counties was necessary as many of the species had multiple introduction points located distant to each other. As did the designing author (Broennimann et al. 2014), I selected the Euclidean distance between the county centroids as my cost function when calculating the MDR. The mean of the error distribution was selected to be 5 years, which provided for temporal uncertainty, but did not produce total randomization (Broennimann et al. 2014). Additionally, I chose to iterate the MDR one hundred times per species output, as it provided an understandable bootstrap value, but did not cause excessive function run time. In summary, the model's selected input parameters were: (1) Error mean (e) =5, (2) Beginning points (r) =2 earliest counties, (3) Iterations (n) =100. The output of the MDR includes for each segment the latitude and longitude of the initiating and recipient nodes, the length of the intervening edge in decimal degrees, the two connecting years with their stochastic subtraction and their difference (dispersal time) , the dispersal rate for that segment, and the associated bootstrap value. The MDR output additionally provides two summary values for the total dispersal distance and median rate.

3.2.5 Application

3.2.5.1 Dispersal Distance

Human mediated dispersal causes more long distance events than purely natural dispersal. To explore the question of whether the species distribution was more characterized by local diffusion or long distance dispersal, I tested whether each species'

estimated observed dispersal distance was closer to the minimum possible or to the distances exhibited by a created randomized distribution. The minimum route length would occur if the expansion pattern approximated pure local diffusion as predicted by Skellam (1951). Therefore, if a species' dispersal distance is closer to the minimum possible than the random, it shows more short distance spread versus longer distance, indicating more localized diffusion. Species which show dispersal distances between minimum and total randomization show a certain amount of longer distance dispersal, but still a level of organized spread.

The species' total network distance is the sum of all the route lengths multiplied by their associated bootstrap value. I computed one route length value for each of the observed and minimum MDRs, and 100 random MDRs were compiled for the null sample distribution. Each randomized network was iterated 100 times for each permutation, for a total of 10,000 cycles for each species. The lengths of the observed routes were compared to both the minimum and the created randomized sample distribution to ascertain whether the species tended toward local or longer distance spread. In order to test whether the species observed route length varied significantly from the random sample distribution a non-parametric Kolmogorov-Smirnov test was performed on each species observed versus random length distribution.

3.2.5.2 Network Connectivity

Human networks tend to display clusters and grow in a non-random fashion. In order investigate the effect of human mediation on the plant's expansion networks I

calculated the number of outgoing edges from each node from all 29 species' networks. By calculating the degree of connectivity within the species networks I could determine if the outgoing edges tended to cluster at certain nodes as seen in scale-free networks. I used the coordinates for the county centroid that initiated the directed graph edge, joined to its Federal Information Processing Standard (FIPS) code, and tallied the number of outgoing edges for each county. The number of edges (*k*) was graphed against the number of nodes which had *k* edges on both a standard and log-log plot to ascertain whether the data approximated a power-law distribution.

To determine if there was a significant difference in the number of edges which originated at a node with a herbarium to one without, I segregated the two groups' edges. As the sample were of different sizes and not transformable to a normal distribution, I performed a Wilcoxon rank-sum test (also known as Mann-Whitney *U* test) to establish the strength of the relationship.

To investigate the influence of herbaria locations and their specimen contributions, I separated from the outgoing edges a subset that only contained initiating nodes collocated with a herbarium. Using the calculated number of outgoing edges, I tested whether the number of specimens provided from the herbarium affected the edge count. Additionally, to distinguish whether herbaria were prominent centers of longer distance dispersal, I mapped the nodes with the highest number of outgoing edges with their long distance dispersal routes, and tested the effect of herbaria specimen contributions on the number of long distance edges. For this analysis I defined *strongly connected* nodes as having a greater quantity of outgoing edges than 2 standard

deviations (SD) above the mean edge count, and *highly connected* to be the nodes with more than 1 SD of edges. Long distance edges were constrained to those which exceeded the length of three average sized Eastern US counties (210 km).

3.2.5.3 Highway Association

To investigate the hypothesis that the species' spread network is associated with human mediation via transportation corridors, I compared the species observed network's coincidence with roads to the level exhibited by the sample random networks. The previously created permuted random and observed networks were shifted from R (R Core Team 2015) to ArcGIS (Esri, Redlands, CA) to create layers of route segments.

Using the weighted, rasterized National Highway Planning Network (NHPN) data layer, I calculated the mean kilometers (km) of highway per square kilometer (km^2) of species' network space (hence referred to as highway-association or HA). I overlaid each species' network segment on the constructed line density NHPN raster and extracted the mean highway-association (HA) value for every segment using the spatial analyst zonal statistic tool. I exported these results as tables and calculated additional summary statistics for each layer. Therefore, each of the 2,929 layers was summarized with an overall mean km/km^2 HA value for its expansion network.

To assess the significance of highway association (HA) I performed a permutation test using the HA values from the random networks to create a sampling distribution.

The permutation number of 100 for each species would allow me to perform a base-line assessment of the species observed networks' HA deviation from random.

3.3 Results

3.3.1 Dispersal Distances

My results showed that two species had observed lengths of their networks that were longer than all instances of the permuted random networks, while a small group of species existed at the opposite extreme with lengths that were near to the minimum possible. The two species with the relatively long networks, *Rosa bracteata* (ROBR) and *Melia azedarach (*MEAZ), are part of the 17% of the species whose lengths of observed networks are closest to being random (Table 3.1). These 5 species, which are within 3 standard deviations of the sample distribution mean, had early specimen collections that were sparse and dispersed widely within their range dimensions. During the early to middle years of their residence time, specimens were collected from isolated counties, with very little contiguity until very close to the present time. *Melia azedarach*, the species with the highest observed network distance, had only 26 counties represented in its distribution during the period from its first specimen in 1885 until 1923. This small number of counties was located spread throughout the boundary encompassing: north to Washington, DC; south to Miami-Dade County, FL; west to Terrebonne Parish, LA; and northwest to St. Louis; MO. Only one specimen was collected outside of these boundaries during the species later residence.

Table 3.1 The species whose observed network lengths are greater than the random distribution or within three standard deviations of the mean. Both *Melia azedarach* (MEAZ) and *Rosa bracteata* (ROBR) had lengths longer than the random routes. Only one species, *Vinca minor* (VIMI2) had a route length that was within the random distribution. Lengths are measured in decimal degrees (D^{o})

Species	Minimum Route Length	Observed Route Length	Mean Random Length	Range of Distribution	Standard Deviation of Distribution
MEAZ	191.50	215.23	210.60	209.01 - 211.70	0.48
ROBR	85.93	11j6	110.66	109.02 - 111.76	0.56
VIMI2	236.42	281.43	280.52	279.19 - 282.09	0.60
MISI	156.14	184.36	186.11	184.88 - 187.91	0.55
NADO	92.88	112.66	114.20	113.21 - 115.65	0.50

On the other side of the dispersal distance continuum there were 10% of the species whose observed network lengths were noticeably closer to the minimum dispersal distance, indicating a lack of long-distance dispersal (Table 3.2). The species *Imperata cylindrica's* (IMCY) observed network distance is almost identical to the minimum route distance possible. Within our overall species group, *I. cylindrica* is the newest plant introduction and has the smallest number of counties represented, with the most compact of all the distributions, focused on the Gulf Coast of the USA. Unlike, *M. azedarach* which was used as an ornamental specimen, *I. cylindrical* was introduced accidentally and has very little practical use except as forage (Dozier et al. 1998). *Microstegium vimineum* (MIVI), as IMCY, has experienced primarily accidental spread to more localized contiguous counties since its accidental introduction in 1919 in the Knoxville, TN area (Fairbrot and Gray 1972). The third species in this group, *Euonymus alata* (EUAL8), has been used as a horticulture species, yet its spread pattern still exhibits regional contiguity. Although qualitative, 5 separate geographical clusters

appear to have emerged from expansion connecting to adjoining counties (Fig. 3.3). This adjacent spread likely produced the shorter length of the species total dispersal distance.

Table 3.2 Species with observed route lengths that are closer to the minimum possible length than to their random distribution. *Imperata cylindrical* (IMCY) was substantially closer to the minimum length than all other species. Lengths are measured in decimal degrees (D^o)

Species	Minimum Route Length	Observed Route length	Mean Random Length	Range of Distribution	Standard Deviation of Distribution
EUAL8	149.67	160.65	175.71	174.32 - 177.08	0.51
MIVI	234.31	247.98	272.37	270.94 - 273.39	0.51
IMCY	92.99	93.92	103.92	102.71 - 104.79	0.38

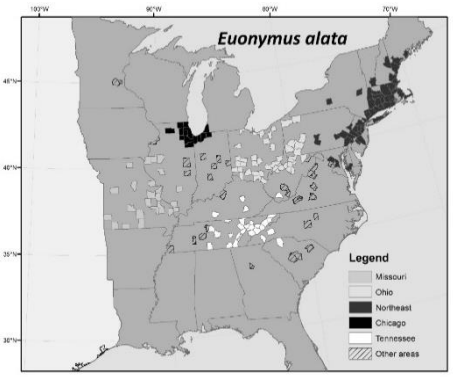


Figure 3.3 Regional distribution of *Euonymus alata's* (EUAL8) which shows distinct geographic areas of clustering. The quantity of contiguity between counties in the different regional areas likely contributed to the shorter distance of its observed network

The remaining 21 species (72%) had lengths of their observed networks that fell moderately between the extremes (Table 3.3). The overall variability in this group showed some species exhibiting spread lengths that approached the sample distribution but still had a structured behavior to their pattern. Of the species in this group, the ones that have less variance between observed and minimum network lengths do tend to exhibit more localized dispersal patterns. Additionally to verify the species'

association with their sample distribution, a Kolmogorov-Smirnov test was performed on all 29 species' observed networks against the random distributions. All but one species showed a significant Kolmogorov-Smirnov test at $p < 0.05$ indicating non-random route lengths, while *Vinca minor* (VIMI2) was verified to be within the random network ($D= 0.94$, $p = 0.139$).

Table 3.3 Species have their observed route length moderately placed between the minimum length and the random distribution lengths. Species are ordered by the ratio of their observed length to the mean of the sample distribution. Lengths are measured in decimal degrees (D^o)

Species	Minimum Route Length	Observed Route Length	Mean Random Length	Range of Distribution	Standard Deviation of Distribution	Ratio of Observed to Random Length
LOTA	198.60	221.52	240.36	238.60 - 242.10	0.643	0.922
LYJA	109.41	124.89	135.34	134.38 - 136.16	0.411	0.923
EUFO5	144.50	159.18	168.78	167.28 - 170.47	0.549	0.943
WISI	165.54	186.43	197.03	195.79 - 198.48	0.597	0.946
ALPE4	306.48	338.02	356.74	355.63 - 359.12	0.656	0.948
LIVU	183.29	209.28	220.55	218.90 - 222.65	0.681	0.949
VIMA	121.11	137.02	143.52	142.37 - 144.35	0.455	0.955
TRSE6	101.26	116.25	121.74	120.46 - 123.02	0.469	0.955
LECU	353.38	382.68	399.92	398.36 - 401.57	0.688	0.957
ELAN	133.63	155.04	161.56	159.71 - 164.13	0.687	0.960
LOMA6	218.31	237.49	246.78	245.39 - 248.00	0.547	0.962
PATO2	191.81	217.95	226.28	225.07 - 227.87	0.532	0.963
LOMO2	209.31	238.15	243.93	242.43 - 246.02	0.679	0.976
ALJU	284.35	321.94	329.44	327.64 - 330.69	0.602	0.977
AIAL	329.28	368.96	377.38	375.70 - 378.72	0.643	0.978
ELUM	291.87	324.56	331.81	330.29 - 333.39	0.704	0.978
LISI	252.86	283.08	289.25	288.10 - 290.41	0.521	0.979
CEOR7	203.26	233.82	238.59	237.22 - 241.08	0.681	0.980
ROMU	399.55	441.35	449.42	447.85 - 451.15	0.739	0.982
LOJA	483.92	537.80	546.34	545.10 - 547.83	0.659	0.984
PUMOL	292.34	333.68	338.81	337.40 - 339.87	0.559	0.985

3.3.2 Network Connectivity

Network classification results indicate that invasive species expansion patterns exhibit qualities of a scale-free network. In scale-free networks there are a large number of nodes that each provide very little connectivity, and a few highly connected nodes that function as hubs. For the aggregated networks of all 29 species, the 2047 eastern counties contained 1662 nodes that had differing numbers of outgoing edges (M=41.76, SD=34.43, N_{edge}=33217). My results show that while most of the represented counties had less than 50 outgoing nodes, a small number exhibited over 150 (Fig. 3.4a). There were 12 nodes that were considered strongly connected, at greater than 2 standard deviations above the mean (110 outgoing edges), with an additional 20 more that were above 1 standard deviation (75 outgoing edges). This type of distribution approximates a scale-free network (Fig. 3.2 power-law curve), although it does not appear as linear (Fig. 3.4b) as larger networks with higher connectivity degrees such as the World Wide Web and movie actor connectivity (Amaral et al. 2000).

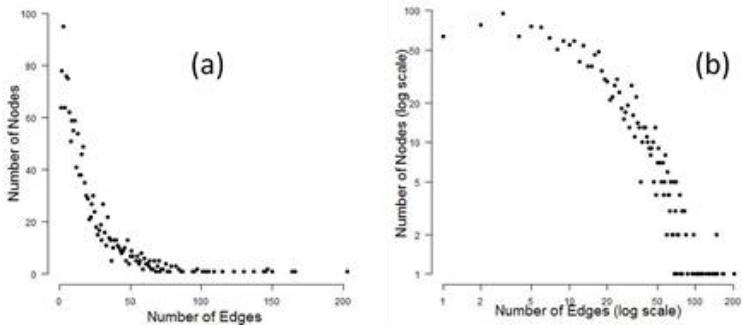


Figure 3.4 (a) The power-law distributions of edges to nodes are long-tailed, with most of the nodes having fewer edges. There are a few nodes which are highly connected (b) The near linear edge distribution of 1662 nodes on a log-log scale. The degree distributions of scale-free networks approximate a power-law

The overall effect of herbaria location indicates that the counties that have herbaria have more outgoing edges than those that do not (Fig. 3.5). To verify the boxplot's visual representation a Wilcoxon rank-sum test with continuity correction (equivalent to the Mann-Whitney *U*-test) for Count of outgoing edges by Herbarium presence indicated that herbaria locations were more likely to have higher outgoing edge counts (*Mdn*=41) than counties without a herbarium (*Mdn*= 13), $W_{(1693)}$=40929, p<0.001.

Using the subset of just the herbaria-occupied counties, results showed that the quantity of specimens provided by the herbarium to my study did not unequivocally explain the number of outgoing edges from that node. Most of the herbaria provided under 500 specimens to the study and exhibit under 100 outgoing edges (Fig. 3.6).

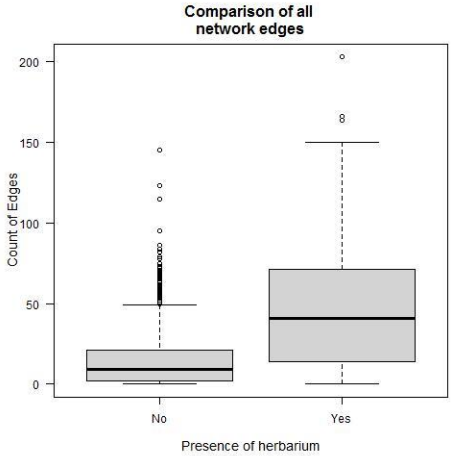


Figure 3.5 Counties (n=101) with a herbarium have significantly higher number of outgoing edges than those without a herbarium (n=1592) (W=40929, p< .001 Mdn(H)=41, Mdn(NH)= 13)

Beyond this degree there are herbaria with high specimen number – low edge counts, and low specimen number – high edge count. While the regression of the number of outgoing edges on total herbarium samples in the study showed significance, it also exhibited only a modest slope ($F_{(1,5314)}$ = 4246, p< 0.001), with r^2 = 0.44. The low slope value (β = 0.028) indicates the effect of increasing the specimen numbers changes

the edge count very little. This implies that among the counties that have a herbarium, the quantity of provided specimens has little influence on their connectivity. Additionally, to investigate the effect of the one outlier, I removed that county from the test and performed the regression again. Without the outlier the results show a slight higher slope, $\beta = 0.045$, but less explained correlation ($F_{(1,5111)} = 2948$, $p < 0.001$, $r^2 = 0.37$). Mapping just the *strongly connected* nodes with their associated edges and the nodes only that represent the *highly connected* edges, reveals noteworthy patterns of species expansions (Fig. 3.7).

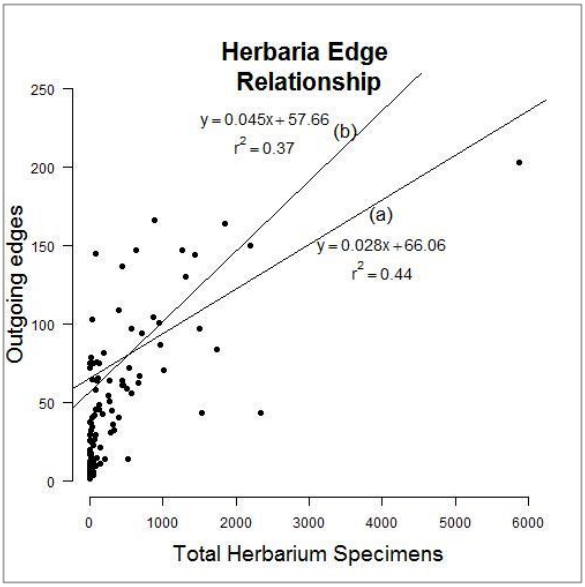


Figure 3.6 Relationship between the number of outgoing edges from a county and the number of herbarium specimens that were provided by a herbarium in that county. (a) Regression using all herbaria show little increase in edge counts as specimen contribution increases. (b) By removing the outlier point (NCU) the regression slope increases, but the amount of effect decreases. Through both evaluations there is little effect on edge count as the number of specimens provided increases. n(counties with herbarium)=101, n(edges)=5316

Even though the group of species comprises different ranges and expansion timing (Appendix C), the longest routes (i.e., > 2 SD) represented 27 of the 29 species, and indicate similarity in invasive spread patterns. The 370 edges and their 12 nodes show distinct well-traveled pathways that also encompass many of the additional 20 *highly connected* nodes (i.e., > 1 SD). The additional placement of the herbaria with their

specimen counts on the map show that it may be the location of the herbaria as opposed to the quantity of specimens that influences the edge number.

When comparing the node locations for both incoming and outgoing directed edges of all lengths, results show that both classes of *highly connected* nodes are generally located in the same geographic area (Fig. 3.8). The nodes were limited to those whose edge numbers were greater than 1 standard deviation (i.e., >1SD (incoming) = 51, >1SD outgoing) =103). While there were more incoming nodes in this class due to the strong hub aspect of the outgoing nodes, they were still aggregated in the same southwesterly trajectory.

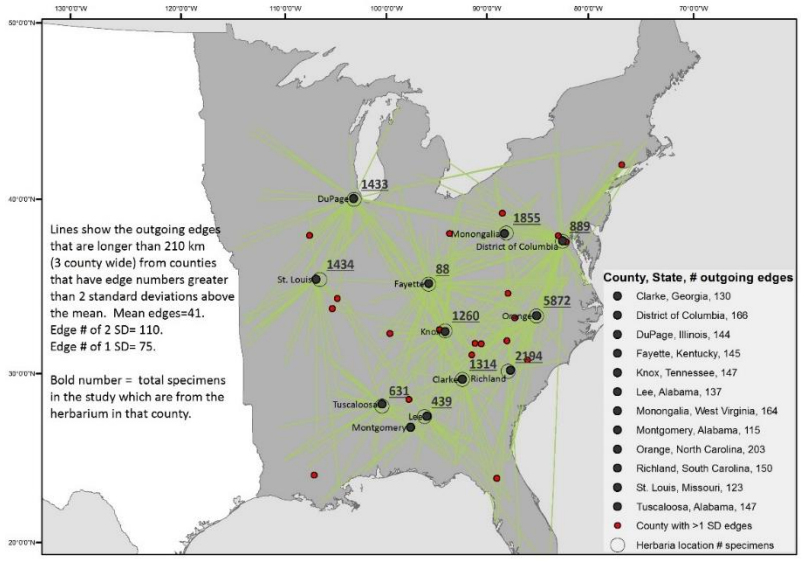


Figure 3.7 Distribution of county locations with one standard deviation (1SD red) and two standard deviations (2SD black) above the average number of edges per county. Herbaria locations for the 2SD nodes and their total contributed specimens. The edges between the 2SD nodes show the areas of highest invasive spread.

The summary analysis of network degree distribution was calculated on the aggregated networks of all 29 species. Most species individually exhibited the same type of connectivity albeit at a smaller scale, although some did show more equality within their degree distribution indicating more random expansion. The aggregated network does show the overlapped routes for all species contained the same common nodes throughout the invasive species distributions regardless of their locations of first occurrences.

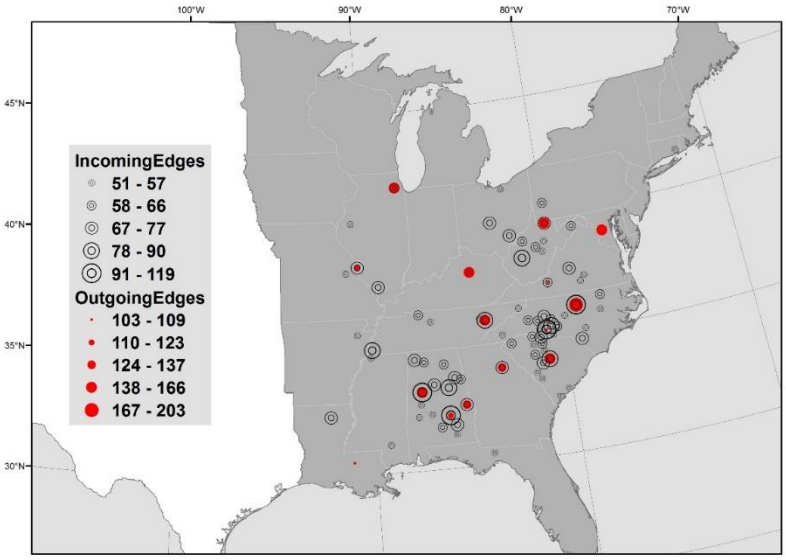


Figure 3.8 The distribution of invasion nodes with high number of incoming and outgoing edges (> mean + 2StDev) Both classes have their largest distribution from states NC to AL.

3.3.3 Highway Association

I found 19 of the 29 (66%) species studied to have their minimum dispersal routes (MDR) more positively associated with highway transportation corridors than would be expected with a purely random dispersal pattern (Fig. 3.9). For these 19 species, the

locations and order of occurrence records indicates that the species spread through a particular pathway that coincided to the locations of the US major road networks. The permutation test allowed me to examine the accuracy of my null hypothesis, which the species were equally likely to spread on routes with or without highway association. The significance level of the association is estimated from the proportion of the sampled distribution's HA that was less than the HA of the observed network. Results of 17 of those 19 species showed mean observed highway-association (HA) to be higher than 100% of the null distribution (p= 0.01). The other two species, *P. tomentosa* and *Wisteria sinensis*, had HA values that were higher than 99% and 93% of the random network's means respectively (Appendix F). Three of the species with the greatest association to the highways also had the largest difference between their observed and random HA mean values. This larger variance could indicate that the time-ordered observed network followed a route structure which predominantly fell over highway routes. While these species, *Ligustrum vulgare*, *Euonymus alata*, and *C. orbiculatus* are also the species with the longest residence time (Appendix A), the trend does not persist. For example, *Microstegium vimineum's* expansion network has the fifth highest coincidence with highways, however, this species is the most recently introduced of the study group.

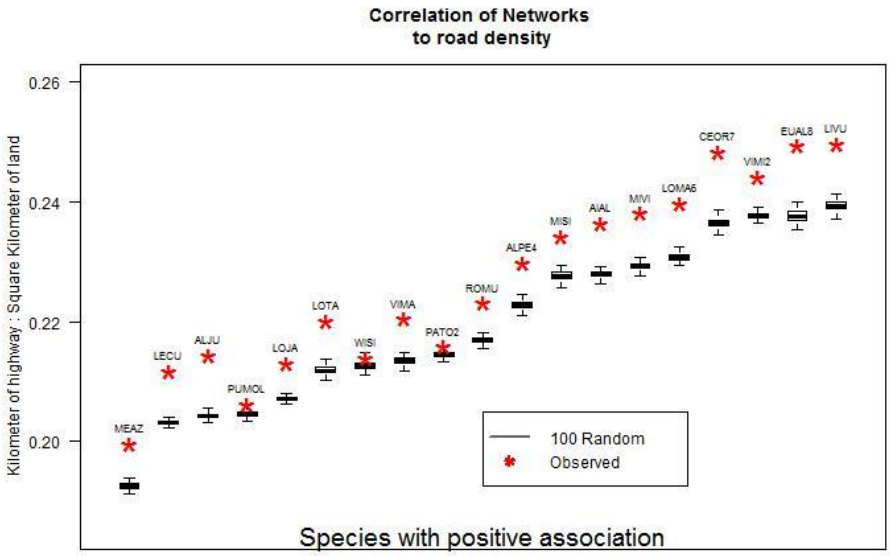


Figure 3.9 Species with higher association between observed networks and highway location compared to random networks

Results also showed a species group with their observed network exhibiting highway association values that were significantly below those expected from random expansion (Fig. 3.10a). Four of the six species to have these negative relationships to highway networks had their observed network HA means below 100% their entire random mean distribution (p= 1). The observed values for the other two species was within the lowest 5% of the null distribution. These lower values than occur by random chance indicate that the species observed routes avoided the areas with major roads during their expansion. Three of the species *R. bracteata* (ROBR) and *Lygodium japonicum* (LYJA) and *Euonymus fortunei* (EUFO5), spread from split distributions in an East-West direction, and additionally E. *fortunei's* herbarium specimens are primarily distributed across the Appalachian Highlands. *Elaeagnus umbellata* (ELUM) which has a much larger distribution than the other species also showed a negative association to the

highway network. This species was highly utilized for mine reclamation and soil stabilization after logging. Therefore, its distribution can be seen to grow out of multiple disconnected clusters centered on more forested areas with connections almost perpendicular to the established highway networks (Fig. 3.10b). Four species exhibited the same highway values as their sample distribution (Fig. 3.11).

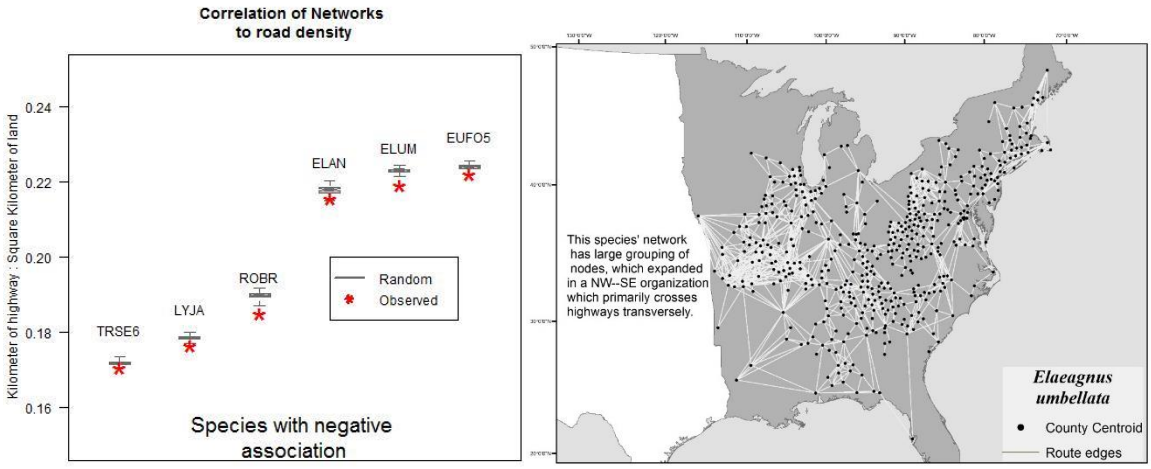


Figure 3.10 (a)Species with lower association between observed networks and highway locations compared to random networks, (b) regional connectivity of *Elaeagnus umbellata* which has high overlap with areas of mine reclamation

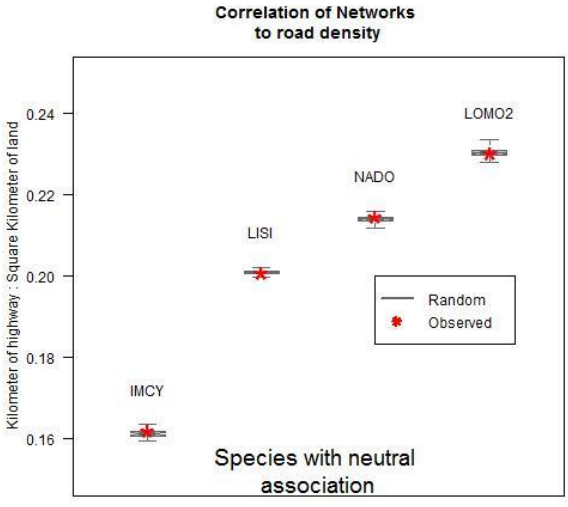


Figure 3.11 Species with similar association between observed networks and highway locations compared to random networks.

3.4 Discussion

Human mediation during a species introduction and invasion can alter natural dispersal patterns causing many measurable changes within the species expansion pathways. We can augment our knowledge regarding the phytogeographic distributions of invasive plants by studying the temporal progression of these spatial patterns. Spatial patterns exhibited during a species expansion may not only be reflective of underlying processes, but may also be contributory to future trajectories of ecologically important processes (Turner 1989). In this study I was able to project the plausible pathways of expansion for 29 long-term invasive plants during the past two centuries. By investigating a large mixed group of plant species (i.e., different traits, life forms, and histories) over a long time period, I could identify exceptional differences in network lengths, connectivity, and association to human transportation corridors.

In regard to dispersal distance, species that spread via human mediated means, rarely spread via only radial, short-distance expansive growth, but also have the addition of longer distance events (Wilson et al. 2009). Within the species group I studied, only one species*, I. cylindrica*, exhibited more local diffusion-type growth than longer-distance expansion. The dispersal distance results also showed that most of the plants had some organization within their expansion network, as shown by the observed distance being less than any of the random dispersal distances. Network distances between minimum and random show the effects of complex interactions with humans. If the species remained in scattered locations only where they were introduced, the

resultant distance would have resembled the random distribution. Subsequent localized spread with location clustering shows the effects of both the human and natural influences. Humans facilitate the alteration of the natural dispersal patterns of alien plants, while also providing the novel environments that allow their establishment and growth.

The locations and connectivity of the nodes reveal an emergence of a scale-free expansion network among this species group. The most strongly connected nodes were found to be collocated with herbaria, although these results may also be an artifact of the colloquial "botanist effect" due to those areas receiving a disproportionate amount of sampling effort (Moerman and Estabrook 2006). Among the herbaria locations, neither the size of the institution, nor the quantity of specimens provided had a large influence on the number of outgoing edges. The critical factor for determining a node's connectivity is its geographical location and placement among human networks. This conclusion follows from the clustering of both the incoming and outgoing directed edges in the same general pathway.

Understanding that invasive species distributions expand approximating scale-free networks has practical utility when prioritizing management options. If a probable invasive species is detected at a less connected node, its chance of large scale invasion is less than if it is introduced into a hub location (Minor and Urban 2008). Scarce resources can then be allocated to a site that is most vulnerable to originate a large scale invasion.

My study explicitly shows that most of the studied species' network connectivity was positively associated with the US highway corridors during the species' expansion. These results build on those of previous studies which concluded that exotic species are aided by transportation corridors to initiate and establish their invasions (Essl et al. 2015; Hulme 2009). Another benefit of this method is the addition of the stochastic component within specimen collection dates, which enabled the study to provide more robust network conclusions. The results therefore showed that the chronological order of specimen records, placed the species networks in line with the current automobile transportation networks. Although many of the earliest specimens were located in counties that were not connected by any major roads, and the US interstate system was not developed until post 1956, the association to these spatial locations allows confidence in this conclusion. Major road construction is created in response to human needs and desires of expedited travel between significant locations (Garrison 1960). Therefore, even though the particular roads may not have been physically constructed at the time of many of my study's network connections, the fact that highways were later instigated in that location, verifies that humans were already traveling on those routes.

Results from all three objectives build on the postulate that the spread of invasive species is inextricably attached to human activities. Analysis of network lengths revealed that the particular chronological order of the species expansion points required spread that was neither strictly diffusional nor random. The patterns additionally were arranged in to a system which had few strong hub locations, such as human constructed

airline route networks. Finally, the majority of the studied species reveal significant positive association with the US highway system. Association to major roads at the regional scale shows that beyond the localized ability to colonize the disturbed edges, the contributory factor of limited access allows a propagule to travel much farther in a limited time. Providing this free way of travel can contribute to much longer distance dispersal and a network pathway that approximates a human construction.

These three findings point to the possibility that much of the problem of invasive plant spread is directly resultant from human mediated dispersal. By combining species biological traits and pathway analysis into future studies, we can gain deeper understanding of the interactions between attributes which enable a species to be more invasive and the routes in which they spread.

This study provides a novel technique to evaluate species distributions and their continuum of differing spatial patterns. Graph theory has proved to be highly useful within many applications from social networks to habitat connectivity. A great attribute of this technique is its flexibility in the additional types of data that can be incorporated. Further work with this dataset can combine landscape coverages, different edge cost functions, or evaluate the results of species' traits versus edge associations. Given the overall global impact of invasive species, the ability to gain deeper understanding of the processes of their spread is essential to future conservation efforts.

REFERENCES

REFERENCES

Ahern RG, Landis DA, Reznicek AA, Schemske DW (2010) Spread of exotic plants in the landscape: the role of time, growth habit, and history of invasiveness. Biological Invasions 12(9):3157-3169

Aikio S, Duncan RP, Hulme PE (2010a) Herbarium records identify the role of long-distance spread in the spatial distribution of alien plants in New Zealand. Journal of Biogeography 37(9):1740-1751

Aikio S, Duncan RP, Hulme PE (2010b) Lag-phases in alien plant invasions: separating the facts from the artefacts. Oikos 119(2):370-378

Aikio S, Duncan RP, Hulme PE (2012) The vulnerability of habitats to plant invasion: disentangling the roles of propagule pressure, time and sampling effort. Global Ecology and Biogeography 21(8):778-786

Amaral LAN, Scala A, Barthelemy M, Stanley HE (2000) Classes of small-world networks. Proceedings of the National Academy of Sciences of the United States of America 97(21):11149-11152

Baker HG (1974) The evolution of weeds. Annual review of ecology and systematics:1-24

Baker HG (1986) Patterns of Plant Invasion in North America. In: Mooney H. andDrake J. (eds), Ecology of Biological Invasions of North America and Hawaii, Ecological Studies. Springer New York, pp. 44-57

Banks NC, Paini DR, Bayliss KL, Hodda M (2014) The role of global trade and transport network topology in the human-mediated dispersal of alien species. Ecology Letters:n/a-n/a

Barabasi AL, Albert R (1999) Emergence of scaling in random networks. Science 286(5439):509-512

Barabasi AL, Bonabeau E (2003) Scale-free networks. Scientific American 288(5):60-69

Barney JN (2006) North American history of two invasive plant species: phytogeographic distribution, dispersal vectors, and multiple introductions. Biological Invasions 8(4):703-717

Barney JN, Whitlow TH, Lembo Jr AJ (2008) Revealing historic invasion patterns and potential invasion sites for two non-native plant species. PLoS One 3(2):e1635

Bauer JT (2012) Invasive species: "back-seat drivers" of ecosystem change? Biological Invasions 14(7):1295-1304

Biggs N, Lloyd EK, Wilson RJ (1976) Graph Theory, 1736-1936. Oxford University Press

Bock F (1971) An algorithm to construct a minimum directed spanning tree in a directed network. Developments in operations research 1(1):29-44

Borger L, Franconi N, De Michele G et al (2006) Effects of sampling regime on the mean and variance of home range size estimates. Journal of Animal Ecology 75(6):1393-1405

Broennimann O, Mraz P, Petitpierre B, Guisan A, Muller-Scharer H (2014) Contrasting spatio-temporal climatic niche dynamics during the eastern and western invasions of spotted knapweed in North America. Journal of Biogeography 41(6):1126-1136

Burrell AM, Pepper AE, Hodnett G et al (2015) Exploring origins, invasion history and genetic diversity of Imperata cylindrica (L.) P. Beauv. (Cogongrass) in the United States using genotyping by sequencing. Molecular Ecology 24(9):2177-2193

Cadotte MW, Murray BR, Lovett-Doust J (2006) Ecological patterns and biological invasions: Using regional species inventories in macroecology. Biological Invasions 8(4):809-821

Capinha C, Essl F, Seebens H, Moser D, Pereira HM (2015) The dispersal of alien species redefines biogeography in the Anthropocene. Science 348(6240):1248-1251

Cavers PB, Heagy MI, Kokron RF (1979) Biology Of Canadian Weeds .35. Alliaria-petiolata (M Bieb) Cavara And Grande. Canadian Journal of Plant Science 59(1):217-229

Charles H, Dukes JS (2007) Impacts of invasive species on ecosystem services. Biological invasions. Springer, pp. 217-237

Chauvel B, Dessaint F, Cardinal-Legrand C, Bretagnolle F (2006) The historical spread of Ambrosia artemisiifolia L. in France from herbarium records. Journal of Biogeography 33(4):665-673

Christen DC, Matlack GR (2009) The habitat and conduit functions of roads in the spread of three invasive plant species. Biological Invasions 11(2):453-465

Crawford PHC, Hoagland BW (2009) Can herbarium records be used to map alien species invasion and native species expansion over the past 100 years? Journal of Biogeography 36(4):651-661

Darwin Core (2004) Darwin Core Schema (version 1.3), A draft standard of the Taxonomic Database Working Group (TDWG)

Delisle F, Lavoie C, Jean M, Lachance D (2003) Reconstructing the spread of invasive plants: taking into account biases associated with herbarium specimens. Journal of Biogeography 30(7):1033-1042

Dobzhansky T (1950) Evolution in the tropics. American Scientist 38(2):209-221

Dozier H, Gaffney JF, McDonald SK, Johnson ERRL, Shilling DG (1998) Cogongrass in the United States: History, Ecology, Impacts, and Management. Weed Technology 12(4):737-743

Durka W, Bossdorf O, Prati D, Auge H (2005) Molecular evidence for multiple introductions of garlic mustard (Alliaria petiolata, Brassicaceae) to North America. Molecular Ecology 14(6):1697-1706

Easley D, Kleinberg J (2010) Networks, crowds, and markets: Reasoning about a highly connected world. Cambridge University Press, UK

Ellis EC (2013) Sustaining biodiversity and people in the world's anthropogenic biomes. Current Opinion in Environmental Sustainability 5(3-4):368-372

Elton CS (1958) The ecology of invasions by animals and plants. University of Chicago Press

Essl F, Bacher S, Blackburn TM et al (2015) Crossing Frontiers in Tackling Pathways of Biological Invasions. Bioscience 65(8):769-782

Fairbrot DE, Gray JR (1972) Microstegium-Vimineum (Trin) A-Camus (Gramineae) in United-States. Bulletin of the Torrey Botanical Club 99(2):97-&

Federal Highway Administration (20140501) In: 14.05 (ed), National Highway Planning Network http://www.fhwa.dot.gov, 14.05. Maks Inc., Columbus, OH,

Fei S, Phillips J, Shouse M (2014) Biogeomorphic Impacts of Invasive Species. In: Futuyma D. J. (ed), Annual Review of Ecology, Evolution, and Systematics, Vol 45, Annual Review of Ecology Evolution and Systematics. pp. 69-+

Ferrari JR, Preisser EL, Fitzpatrick MC (2014) Modeling the spread of invasive species using dynamic network models. Biological Invasions 16(4):949-960

Flannery MCfse (2013) Plant Collections Online: Using Digital Herbaria in Biology Teaching. Bioscene 39(1):3-9

Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME (2005) The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proceedings of the National Academy of Sciences of the United States of America 102(27):9673-9678

Fuentes N, Ugarte E, Kuehn I, Klotz S (2008) Alien plants in Chile: inferring invasion periods from herbarium records. Biological Invasions 10(5):649-657

Garrison WL (1960) Connectivity of the interstate highway system. Papers in Regional Science 6(1):121-137

Gelbard JL, Belnap J (2003) Roads as conduits for exotic plant invasions in a semiarid landscape. Conservation Biology 17(2):420-432

Gonzalez MC, Hidalgo CA, Barabasi A-L (2008) Understanding individual human mobility patterns. Nature 453(7196):779-782

Graham CH, Ferrier S, Huettman F, Moritz C, Peterson AT (2004) New developments in museum-based informatics and applications in biodiversity analysis. Trends in Ecology & Evolution 19(9):497-503

Hauff C, Nürnberger A (2005) On the Use of Scale-Free Networks for Information Network Modelling. Proc. of 1st European Symposium on Nature-inspired Smart Information Systems.

Heffernan JB, Soranno PA, Angilletta MJ et al (2014) Macrosystems ecology: understanding ecological patterns and processes at continental scales. Frontiers in Ecology and the Environment 12(1):5-14

Heywood VH (1989) Patterns, extents and modes of invasions by terrestrial plants. Biological invasions: a global perspective:31-60

Hoekman J, Frenken K, Tijssen RJW (2010) Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe. Research Policy 39(5):662-673

Hordijk W, Broennimann O (2012) Dispersal routes reconstruction and the minimum cost arborescence problem. Journal of Theoretical Biology 308:115-122

Hulme PE (2009) Trade, transport and trouble: managing invasive species pathways in an era of globalization. Journal of Applied Ecology 46(1):10-18

Hulme PE, Bacher S, Kenis M et al (2008) Grasping at the routes of biological invasions: a framework for integrating pathways into policy. Journal of Applied Ecology 45(2):403-414

Kannan R, Shackleton CM, Shaanker RU (2013) Reconstructing the history of introduction and spread of the invasive species, Lantana, at three spatial scales in India. Biological Invasions 15(6):1287-1302

Koh I, Rowe HI, Holland JD (2013) Graph and circuit theory connectivity models of conservation biological control agents. Ecological Applications 23(7):1554-1573

Larkin DJ (2012) Lengths and correlates of lag phases in upper-Midwest plant invasions. Biological Invasions 14(4):827-838

Lavoie C (2013) Biological collections in an ever changing world: Herbaria as tools for biogeographical and environmental studies. Perspectives in Plant Ecology Evolution and Systematics 15(1):68-76

Lavoie C, Jodoin Y, De Merlis AG (2007) How did common ragweed (Ambrosia artemisiifolia L.) spread in Québec? A historical analysis using herbarium records. Journal of Biogeography 34(10):1751-1761

Lelong B, Lavoie C, Jodoin Y, Belzile F (2007) Expansion pathways of the exotic common reed (Phragmites australis): a historical and genetic analysis. Diversity and Distributions 13(4):430-437

Lemke D, Hulme PE, Brown JA, Tadesse W (2011) Distribution modelling of Japanese honeysuckle (Lonicera japonica) invasion in the Cumberland Plateau and Mountain Region, USA. Forest Ecology and Management 262(2):139-149

Levin SA (1992) The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. Ecology 73(6):1943-1967

Liang Y (2010) Exotic invasive plants in Kentucky. Master's Thesis, University of Kentucky

Liebhold AM, McCullough DG, Blackburn LM, Frankel SJ, Von Holle B, Aukema JE (2013) A highly aggregated geographical distribution of forest pest invasions in the USA. Diversity and Distributions 19(9):1208-1216

Loiselle BA, Jorgensen PM, Consiglio T et al (2008) Predicting species distributions from herbarium collections: does climate bias in collection sampling influence model outcomes? Journal of Biogeography 35(1):105-116

Lonsdale WM (1999) Global patterns of plant invasions and the concept of invasibility. Ecology 80(5):1522-1536

Lyons DA, Scheibling RE (2009) Range expansion by invasive marine algae: rates and patterns of spread at a regional scale. Diversity and Distributions 15(5):762-775

Martin SL, Jasinski BL, Kendall AD, Dahl TA, Hyndman DW (2015) Quantifying beaver dam dynamics and sediment retention using aerial imagery, habitat characteristics, and economic drivers. Landscape Ecology 30(6):1129-1144

Matthew WD (1914) Climate And Evolution. Annals of the New York Academy of Sciences 24(1):171-318

McGarigal K, Cushman SA, Neel MC, Ene E (2012) FRAGSTATS: Spatial Pattern Analysis Program for Categorical and Continuous Maps (v4.2.1.603). Computer software program produced by the authors at the University of Massachuessets, Amherst,

Meyer C, Kreft H, Guralnick R, Jetz W (2015) Global priorities for an effective information basis of biodiversity distributions. Nature communications 6:8221-8221

Mihulka S, Pyšek P (2001) Invasion history of Oenothera congeners in Europe: a comparative study of spreading rates in the last 200 years. Journal of Biogeography 28(5):597-609

Milbau A, Stout JC, Graae BJ, Nijs I (2009) A hierarchical framework for integrating invasibility experiments incorporating different factors and spatial scales. Biological Invasions 11(4):941-950

Millennium Assessment Board (2005) Millennium ecosystem assessment. Washington, DC: New Island.

Miller JR, Turner MG, Smithwick EAH, Dent CL, Stanley EH (2004) Spatial Extrapolation: The Science of Predicting Ecological Patterns and Processes. BioScience 54(4):310-320

Minor ES, Urban DL (2008) A graph-theory frarmework for evaluating landscape connectivity and conservation planning. Conservation Biology 22(2):297-307

Moerman DE, Estabrook GF (2006) The botanist effect: counties with maximal species richness tend to be home to universities and botanists. Journal of Biogeography 33(11):1969-1974

Mooney HA, Cleland EE (2001) The evolutionary impact of invasive species. Proceedings of the National Academy of Sciences of the United States of America 98(10):5446-5451

Moravcová L, Pyšek P, Jarošík V, Pergl J (2015) Getting the Right Traits: Reproductive and Dispersal Characteristics Predict the Invasiveness of Herbaceous Plant Species. PLoS ONE 10(4):e0123634

Muggeo VMR (2008) segmented:an R Package to Fit Regression Models with Broken-Line Relationships. R News. pp. 20-25

Munger GT (2001) Alliaria petiolata. , Fire Effects Information System, [Online]. Fire Sciences Laboratory (Producer), Rocky Mountain Research Station,,

Nuzzo V (1993) Distribution and Spread of the Invasive Biennial *Alliaria petiolata* (garlic mustard) in North-America. In: McKnight B. N. (ed), Biological Pollution: The Control and Impact of Invasive Exotic Species. Indiana Academy of Sciences, Indianapolis, pp. 137-145

Nuzzo V (1999) Invasion pattern of herb garlic mustard (Alliaria petiolata) in high quality forests. Biological Invasions 1(2/3):169-179

Oswalt SN, Oswalt CM The extent of selected nonnative invasive plants on southern forest lands. In: Fei S., Lhotka J. M., Stringer J. W., Gottschalk K. W., Miller G. W. (eds) 17th central hardwood forest conference, Lexington, KY 2011. U.S. Department of Agriculture, Forest Service, Northern Research Station, p. 447-459

Peters DPC, Bestelmeyer BT, Turner MG (2007) Cross-scale interactions and changing pattern-process relationships: Consequences for system dynamics. Ecosystems 10(5):790-796

Pianka ER (1966) Latitudinal Gradients In Species Diversity - A Review Of Concepts. American Naturalist 100(910):33-&

Pielou EC (1979) Biogeography. Wiley, New York

Pimentel D, Zuniga R, Morrison D (2005) Update on the environmental and economic costs associated with alien-invasive species in the United States. Ecological economics 52(3):273-288

Powell KI, Chase JM, Knight TM (2013) Invasive Plants Have Scale-Dependent Effects on Diversity by Altering Species-Area Relationships. Science 339(6117):316-318

Pyke GH, Ehrlich PR (2010) Biological collections and ecological/environmental research: a review, some observations and a look to the future. Biological Reviews 85(2):247-266

Pysek P (1991) Heracleum-Mantegazzianum In The Czech Republic - Dynamics Of Spreading From The Historical-Perspective. Folia Geobotanica & Phytotaxonomica 26(4):439-454

Pysek P, Hulme PE (2005) Spatio-temporal dynamics of plant invasions: Linking pattern to process. Ecoscience 12(3):302-315

Pyšek P, Jarošík V, Hulme PE et al (2012) A global assessment of invasive plant impacts on resident species, communities and ecosystems: the interaction of impact measures, invading species' traits and environment. Global Change Biology 18(5):1725-1737

Pysek P, Prach K (1993) Plant Invasions And The Role Of Riparian Habitats - A Comparison Of 4 Species Alien To Central-Europe. Journal of Biogeography 20(4):413-420

Pysek P, Prach K (1995) Invasion dynamics of *Impatiens-glandulifera* - A Century Of Spreading Reconstructed. Biological Conservation 74(1):41-48

R Core Team (2015) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria,

Radford AE, Ahles HE, Bell CR (1968) Manual of the vascular flora of the Carolinas. Univ of North Carolina Press, Chapel Hill

Radford AE, Dickison WC, Massey JR, Bell CR (1974) Vascular plant systematics. Ch. 25, The HERBARIUM. Harper & Row, New York,

Radford LS (1984) The History of the Herbarium of the University of North Carolina at Chapel Hill, NC, 1908-1998. UNC Herbarium, Chapel Hill

Rangel TFLVB, Felizola Diniz-Filho JA, Bini LM (2006) Towards an integrated computational tool for spatial analysis in macroecology and biogeography. Global Ecology and Biogeography 15(4):321-327

Reichard SH, White P (2001) Horticulture as a pathway of invasive plant introductions in the United States. Bioscience 51(2):103-113

Ricciardi A (2007) Are modern biological invasions an unprecedented form of global change? Conservation Biology 21(2):329-336

Robertson T, Döring M, Guralnick R et al (2014) The GBIF Integrated Publishing Toolkit: Facilitating the Efficient Publishing of Biodiversity Data on the Internet. PLoS ONE 9(8):e102623

Rouget M, Richardson David M (2003) Inferring Process from Pattern in Plant Invasions: A Semimechanistic Model Incorporating Propagule Pressure and Environmental Factors. The American Naturalist 162(6):713-724

Roush W, Munroe JS, Fagre DB (2007) Development of a spatial analysis metho using ground-based repeat photography to detect changes in the alpine treeline ecotone, Glacier National Park, Montana, USA. Arctic Antarctic and Alpine Research 39(2):297-308

Schierenbeck KA (2004) Japanese honeysuckle (Lonicera japonica) as an invasive species; history, ecology, and context. Critical reviews in plant sciences 23(5):391-400

Shigesada N, Kawasaki K (1997) Biological invasions: theory and practice. Oxford University press, UK

Shigesada N, Kawasaki K, Takeda Y (1995) Modeling stratified diffusion in biological invasions. American Naturalist 146(2):229-251

Simberloff D (2011) How common are invasion-induced ecosystem impacts? Biological Invasions 13(5):1255-1268

Simberloff D, Schmitz DC, Brown TC (1997) Strangers in paradise: impact and management of nonindigenous species in Florida. Island press, Washington, D.C.

Skellam JG (1951) Random Dispersal In Theoretical Populations. Biometrika 38(1-2):196-218

Soranno PA, Cheruvelil KS, Bissell EG et al (2014) Cross-scale interactions: quantifying multiscaled cause-effect relationships in macrosystems. Frontiers in Ecology and the Environment 12(1):65-73

Strayer DL, Eviner VT, Jeschke JM, Pace ML (2006) Understanding the long-term effects of species invasions. Trends in Ecology & Evolution 21(11):645-651

Suarez AV, Holway DA, Case TJ (2001) Patterns of spread in biological invasions dominated by long-distance jump dispersal: insights from Argentine ants. Proceedings of the National Academy of Sciences 98(3):1095-1100

Suarez AV, Tsutsui ND (2004) The Value of Museum Collections for Research and Society. BioScience 54(1):66-74

Thiers BM (continuously updated) Index Herbariorum: A global directory of public herbaria and associated staff. New York Botanical Garden's Virtual Herbarium. Available from http://sweetgum.nybg.org/science/ih/

Turner MG (1989) Landscape Ecology - The Effect of Pattern on Process. Annual Review of Ecology and Systematics 20:171-197

US Forest Service (2005) Factsheets: Garlic Mustard (*Alliaria petiolata*). St. Paul, MN. Available from http://www.na.fs.fed.us/fhp/invasive_plants/weeds/garlic_mustard.pdf

van Kleunen M, Dawson W, Maurel N (2015) Characteristics of successful alien plants. Molecular Ecology 24(9):1954-1968

van Kleunen M, Dawson W, Schlaepfer D, Jeschke JM, Fischer M (2010) Are invaders different? A conceptual framework of comparative approaches for assessing determinants of invasiveness. Ecology Letters 13(8):947-958

Vitousek PM, Dantonio CM, Loope LL, Rejmanek M, Westbrooks R (1997) Introduced species: A significant component of human-caused global change. New Zealand Journal of Ecology 21(1):1-16

Weber E (1998) The dynamics of plant invasions: a case study of three exotic goldenrod species (Solidago L.) in Europe. Journal of Biogeography 25(1):147-154

White EP, Enquist BJ, Green JL (2008) On Estimating the Exponent of Power-Law Frequency Distributions. Ecology 89(4):905-912

Wilkinson DM (2001) Dispersal: Biogeography. eLS. John Wiley & Sons, Ltd,

Williamson M, Pysek P, Jarosik V, Prach K (2005) On the rates and patterns of spread of alien plants in the Czech Republic, Britain, and Ireland. Ecoscience 12(3):424-433

Wilson JRU, Dormontt EE, Prentis PJ, Lowe AJ, Richardson DM (2009) Something in the way you move: dispersal pathways affect invasion success. Trends in Ecology & Evolution 24(3):136-144

With KA (2002) The landscape ecology of invasive spread. Conservation Biology 16(5):1192-1203

Wu J, Jenerette GD, Buyantuyev A, Redman CL (2011) Quantifying spatiotemporal patterns of urbanization: The case of the two fastest growing metropolitan regions in the United States. Ecological Complexity 8(1):1-8

APPENDICES

Appendix A Study Species

Appendix A 29 Species used in study

Code	Common Name	Family	Genus	Species	CtY #	MRT	First Spec	Orign	N Fix	Life Form	Blom	Tiss	Disp	Why	Hort
AIAL	Tree-of-Heaven	Simaroubaceae	*Ailanthus*	*altissima*	654	173	1841	Asia	no	TREE	mid	woody	Wind	Use	No
ALJU	Silktree, Mimosa	Fabaceae	*Albizia*	*julibrissin*	532	143	1871	Asia	yes	TREE	mid	woody	Animal	Hort	Yes
ALPE4	Garlic Mustard	Brassicaceae	*Alliaria*	*petiolata*	581	146	1868	Europe	no	FORB	early	herb	Other	Acc	No
CEOR7	Oriental Bittersweet	Celastraceae	*Celastrus*	*orbiculatus*	298	174	1840	Asia	no	VINE	mid	woody	Animal	Hort	Yes
ELAN	Russian-olive	Elaeagnaceae	*Elaeagnus*	*angustifolia*	138	141	1873	Europe	yes	TREE	mid	woody	Animal	Use	No
ELUM	Autumn-Olive	Elaeagnaceae	*Elaeagnus*	*umbellata*	484	141	1873	Asia	yes	SHRUB	early	woody	Animal	Use	No
EUAL8	Winged burning bush	Celastraceae	*Euonymus*	*alata*	203	128	1886	Asia	no	SHRUB	mid	woody	Animal	Hort	Yes
EUFO5	Winter creeper	Celastraceae	*Euonymus*	*fortunei*	202	126	1888	Asia	no	VINE	mid	woody	Animal	Hort	Yes
IMCY	Cogongrass	Poaceae	*Imperata*	*cylindrica*	129	91	1923	Asia	no	GRASS	early	herb	Wind	Use	No
LECU	Chinese lespedeza	Fabaceae	*Lespedeza*	*cuneata*	697	86	1928	Asia	yes	FORB	late	woody	Animal	Use	No
LISI	Chinese privet	Oleaceae	*Ligustrum*	*sinense*	483	110	1904	Asia	no	SHRUB	mid	wood	Anim	Hort	Yes
LIVU	European privet	Oleaceae	*Ligustrum*	*vulgare*	249	192	1822	Europe	no	SHRUB	mid	woody	Anim	Hort	Yes
LOJA	Japanese honeysuckle	Caprifoliaceae	*Lonicera*	*japonica*	1075	146	1868	Asia	no	VINE	mid	wood	Anim	Hort	Yes
LOMA6	Amur honeysuckle	Caprifoliaceae	*Lonicera*	*maackii*	367	114	1900	Asia	yes	SHRUB	early	wood	Anim	Use	No
LOMO2	Morrows honeysuckle	Caprifoliaceae	*Lonicera*	*morrowii*	334	115	1899	Asia	no	SHRUB	early	wood	Anim	Hort	Yes
LOTA	Tatarian honeysuckle	Caprifoliaceae	*Lonicera*	*tatarica*	261	141	1873	Europe	no	SHRUB	early	wood	Anim	Hort	Yes

LYJA	Japanese climbing fern	Lygodiaceae	*Lygodium*	*japonicum*	211	133	1881	Asia	no	VINE	mid	herb	Wind	Hort	Yes
MEAZ	Chinaberry Tree	Meliaceae	*Melia*	*azedarach*	340	129	1885	Asia	no	TREE	mid	wood	Anim	Hort	Yes
MISI	Chinese silvergrass	Poaceae	*Miscanthus*	*sinensis*	215	165	1849	Asia	no	GRASS	late	herb	Wind	Hort	Yes
MIVI	Japanese Stiltgrass,	Poaceae	*Microstegium*	*vimineum*	475	83	1931	Asia	no	GRASS	late	herb	Wind	Acc	No
NADO	Sacred Bamboo	Berberidaceae	*Nandina*	*domestica*	115	135	1879	Asia	no	SHRUB	mid	wood	Anim	Hort	Yes
PATO2	Princesstree	Scrophulariaceae	*Paulownia*	*tomentosa*	325	142	1872	Asia	no	TREE	early	wood	Wind	Hort	Yes
PUMOL	Kudzu	Fabaceae	*Pueraria*	*montana var. lobata*	537	114	1900	Asia	yes	VINE	late	wood	Other	Use	No
ROBR	McCartney Rose	Rosaceae	*Rosa*	*bracteata*	122	146	1868	Asia	no	SHRUB	early	wood	Anim	Hort	No
ROMU	Multiflora rose	Rosaceae	*Rosa*	*multiflora*	773	118	1896	Asia	no	SHRUB	mid	wood	Anim	Hort	Yes
TRSE6	Tallowtree	Euphorbiaceae	*Triadica*	*sebifera*	151	163	1851	Asia	no	TREE	mid	wood	Anim	Use	No
VIMA	Bigleaf periwinkle	Apocynaceae	*Vinca*	*major*	152	145	1869	Europe	no	VINE	early	herb	Other	Hort	Yes
VIMI2	Common periwinkle	Apocynaceae	*Vinca*	*minor*	402	192	1822	Europe	no	VINE	early	herb	Other	Hort	Yes
WISI	Chinese wisteria	Fabaceae	*Wisteria*	*sinensis*	240	135	1879	Asia	yes	VINE	early	wood	Other	Hort	Yes

Appendix B Herbaria

Appendix B Herbaria list

CODE	ID	Count
AMAL	Anniston Museum of Natural History	35
ANSP	Academy of Natural Sciences	60
APSC	Austin Peay State University	256
ARCH	Archbold Biological Station	11
ARIZ	University of Arizona	206
ASC	Northern Arizona University	150
ASU	Arizona State University	182
AUA	Auburn University	439
BALT	Towson University	79
BAYLU	Baylor University	1
BEREA	Berea College	303
BH	Cornell University	682
BHO	Ohio University	131
BLH	Cranbrook Institute of Science	3
BRY	Brigham Young University	28
BUF	Buffalo Museum of Science	11
BUPL	Bucknell University	59
BUT	Butler University	90
CAS	California Academy of Sciences	1
CCH	University of Arizona South Agricultural Extension Service	10
CEL	University of Illinois	54
CHIC	Chicago Botanic Garden	58
CHRB	Rutgers University	282
CLEMS	Clemson University	1743
CM	Carnegie Museum of Natural History	1501
CONN	University of Connecticut	2336
CS	Colorado State University	41
DEK	Northern Illinois University	7
DES	Desert Botanical Garden	118

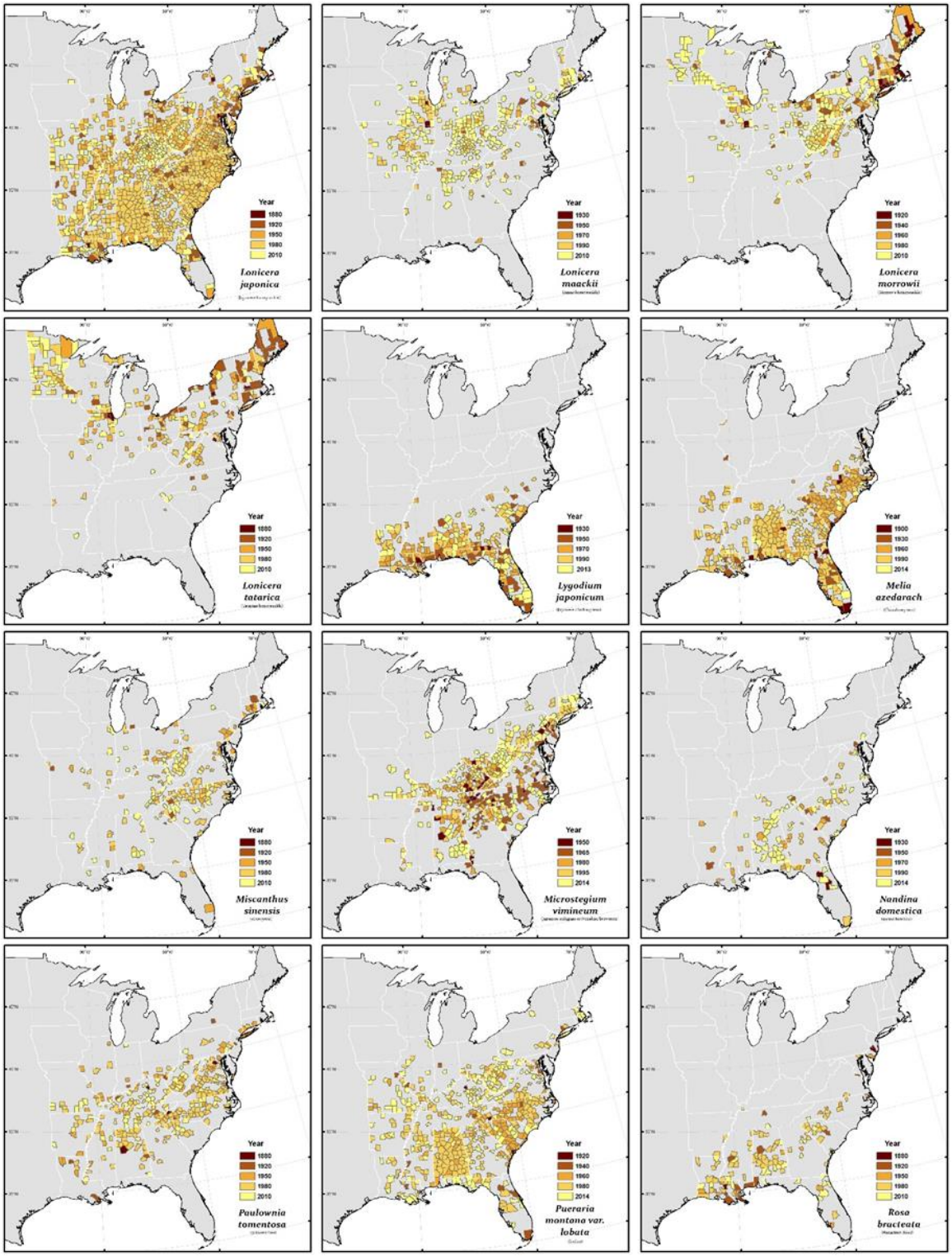
DSC	Delta State University	13
DUKE	Duke University	661
EIU	Eastern Illinois University	17
EKY	Eastern Kentucky University	401
EPHR	Snow College	14
F	Field Museum of Natural History	130
FARM	Longwood University	24
FLAS	Florida Museum of Natural History	965
FLD	Fort Lewis College	15
FSU	Florida State University	535
FTG	Fairchild Tropical Botanic Garden	142
FUGR	Furman University	1
GA	University of Georgia	1314
GAS	Georgia Southern University	325
GCNP	Grand Canyon National Park	16
GEO	Emory University	15
GH	Harvard University	502
GMUF	George Mason University	54
HNH	Dartmouth College	4
HPC	Howard Payne University	200
HTTU	Tennessee Technological University	27
HWBA	Benedictine College	2
IBE	Institute for Botanical Exploration	89
ID	University of Idaho	2
ILL	University of Illinois	16
ILLS	Illinois Natural History Survey	1006
IND	Indiana University	7
ISC	Iowa State University	46
ISM	Illinois State Museum	37
ISTC	University of Northern Iowa	16
JEF	Indiana University Southeast	73
JSU	Jacksonville State University	506
KANU	University of Kansas	1
KE	Kent State University	27
KHD	Denver Botanic Gardens	68
KNK	Northern Kentucky University	28
KSC	Kansas State University	225

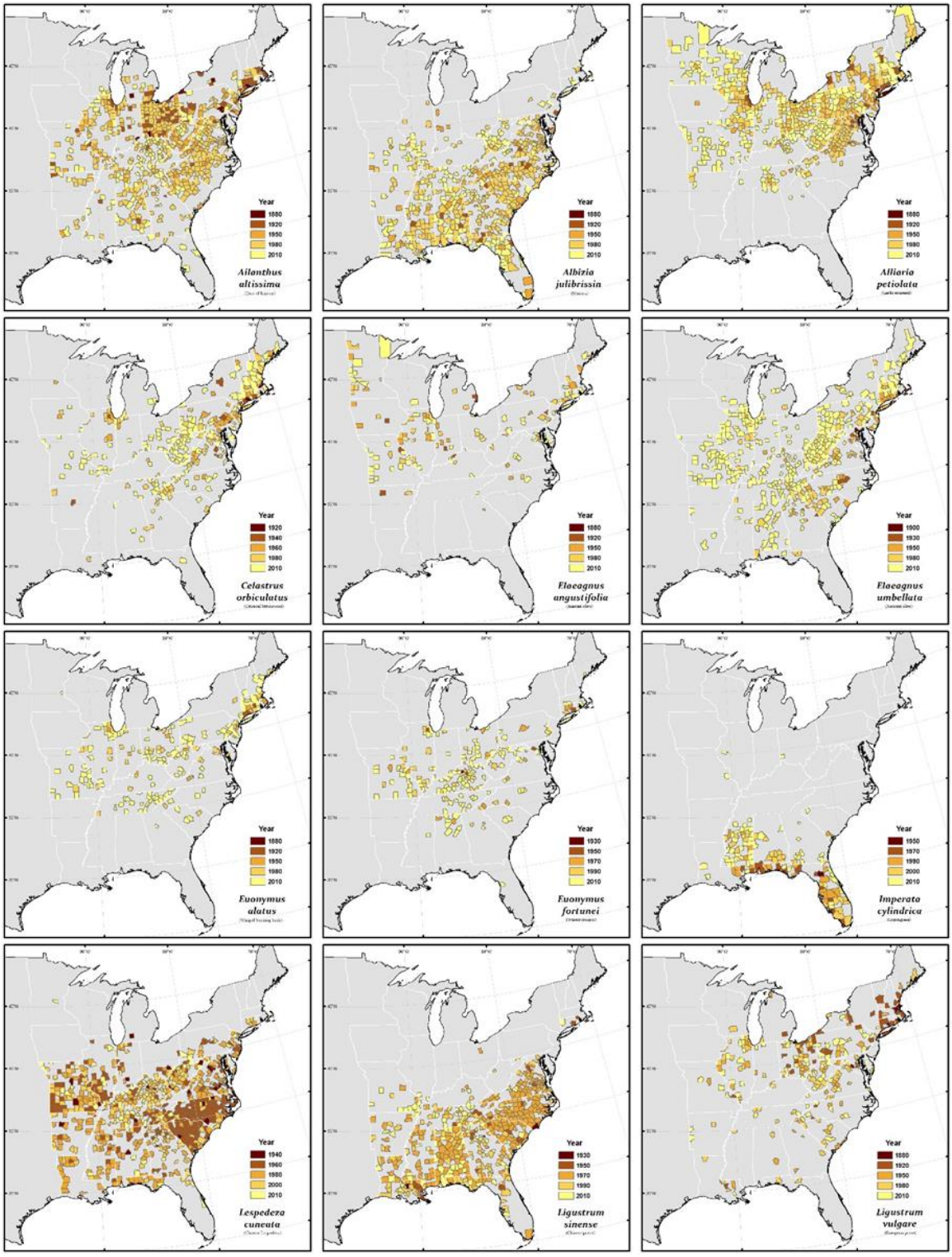
KSP	Pittsburg State University	7
KSTC	Emporia State University	8
LCDI	Luther College	3
LEH	Lehigh University	6
LSU	Louisiana State University	869
LYN	Lynchburg College	406
MAINE	University of Maine	78
MARY	University of Maryland	195
MASS	University of Massachusetts	21
MDKY	Morehead State University	145
MICH	University of Michigan	8
MIL	Milwaukee Public Museum	30
MIN	University of Minnesota	215
MISS	University of Mississippi	443
MISSA	Mississippi State University	74
MISU	Minot State University	36
MNA	Museum of Northern Arizona	31
MO	Missouri Botanical Garden	1434
MOAR	Morris Arboretum University of Pennsylvania	232
MODNR	Division of State Parks	22
MOR	Morton Arboretum	1433
MSC	Michigan State University	62
MU	Miami University	332
MUR	Murray State University	109
MVSC	Millersville University	21
MWI	Western Illinois University	709
NAVA	Navajo Nation Department of Fish and Wildlife	19
NCSC	North Carolina State University	103
NCU	University of North Carolina	5872
ND	University of Notre Dame	23
NDG	University of Notre Dame	4
NEBK	University of Nebraska at Kearney	1
NEMO	Truman State University	27
NHA	University of New Hampshire	12
NMC	New Mexico State University	40

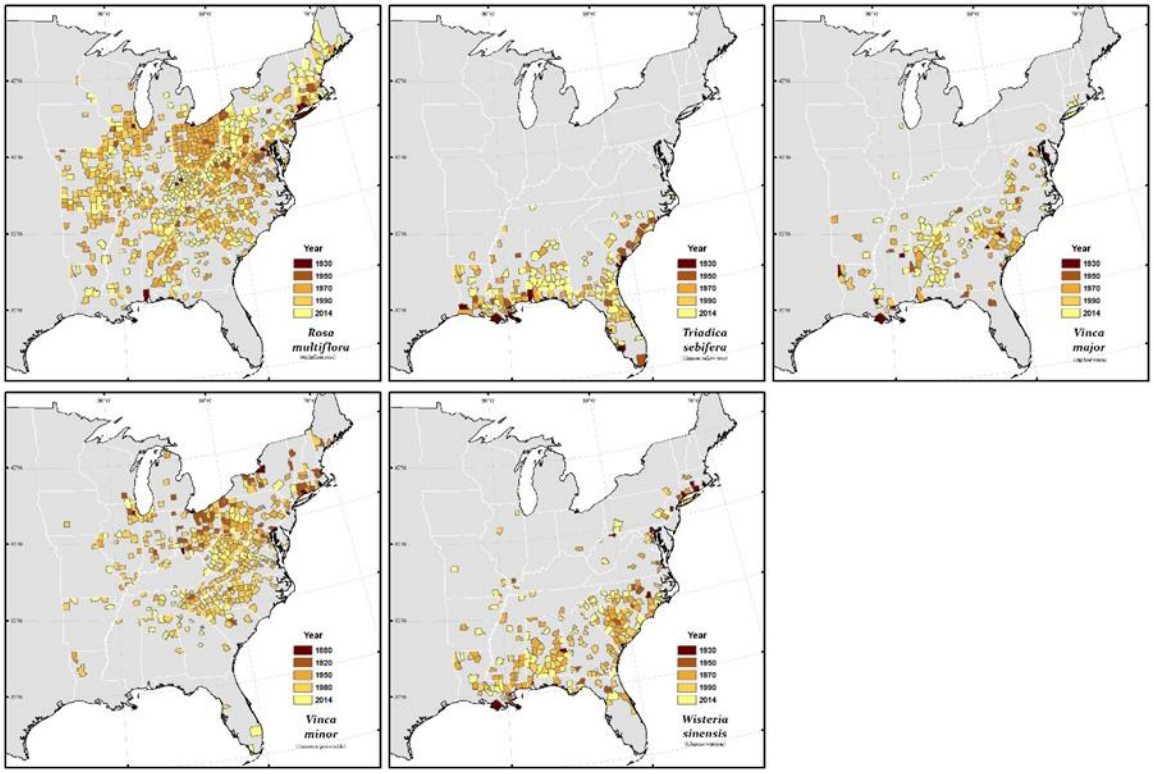
NMCR	New Mexico State University	8
NO	Tulane University	268
NY	New York Botanical Garden	1528
NYS	New York State Museum	27
ODU	Old Dominion University	37
OS	Ohio State University	949
OSC	Oregon State University	10
OSH	University of Wisconsin Oshkosh	12
PAC	The Pennsylvania State University	23
PAM	Pennsylvania Department of Agriculture	9
PUL	Purdue University	2
PUR	Purdue University	95
RM	University of Wyoming	300
SAMF	Samford University	10
SC	Salem College	26
SDC	South Dakota State University	3
SDSU	San Diego State University	40
SEMO	Southeast Missouri State University	96
SJNM	San Juan College	112
SMS	Missouri State University	45
SNM	Western New Mexico University	39
SSLP	Rocky Mountain Research Station	3
STAR	Arkansas State University	125
SUU	Southern Utah University	3
SYRF	State University of New York	3
TAC	Tarleton State University	1
TENN	University of Tennessee	1260
TEUI	U.S. Forest Service Southwest Region Terrestrial Ecological Unit Inventory	19
TROY	Troy University	438
TTRS	Tall Timbers Research Station	215
UAM	University of Arkansas at Monticello	19
UARK	University of Arkansas	5
UCR	University of California	477
UMO	University of Missouri	280

UNA	University of Alabama	631
UNCC	Mecklenburg County Park and Recreation Herbarium	33
UNM	University of New Mexico	109
US	Smithsonian Institution	889
USAM	University of South Alabama	6
USCH	University of South Carolina	2194
USF	University of South Florida	519
USUUB	Utah State University Uintah Basin	3
UTC	Utah State University	99
UVSC	Utah Valley University	45
UWAL	University of West Alabama	182
UWGB	University of Wisconsin-Green Bay	5
UWL	University of Wisconsin	3
UWM	University of Wisconsin	22
UWSP	University of Wisconsin-Stevens Point	47
VMI	Virginia Military Institute	111
VPI	Virginia Polytechnic Institute and State University	396
VSC	Valdosta State University	461
VT	University of Vermont	7
WARM	University of Central Missouri	1
WCUH	Western Carolina University	578
WIS	University of Wisconsin	106
WKU	Western Kentucky University	54
WSC	Western State College	2
WSFA	Wilmington College	1
WTU	University of Washington	24
WVA	West Virginia University	1855
WVW	West Virginia Wesleyan College	121
YU	Yale University	569
	TOTAL COUNT--	44991

Appendix C Species spatiotemporal distributions




Year
1880
1920
1950
1980
2010
Ailanthus altissima
Albizia julibrissin
Alliario petiolata
Year
1920
1940
1960
1980
2010
Celastrus orbiculatus
Year
1880
1920
1950
1980
2010
Elaeagnus angustifolia
Year
1900
1930
1950
1980
2010
Elaeagnus umbellata
Year
1880
1920
1950
1980
2010
Euonymus alatus
Year
1930
1950
1970
1990
2010
Euonymus fortunei
Year
1950
1970
1990
2000
2010
Imperata cylindrica
Year
1940
1960
1980
2000
2010
Lespedeza cuneata
Year
1930
1950
1970
1990
2010
Ligustrum sinense
Year
1880
1920
1950
1980
2010
Ligustrum vulgare



Appendix C Species distributions

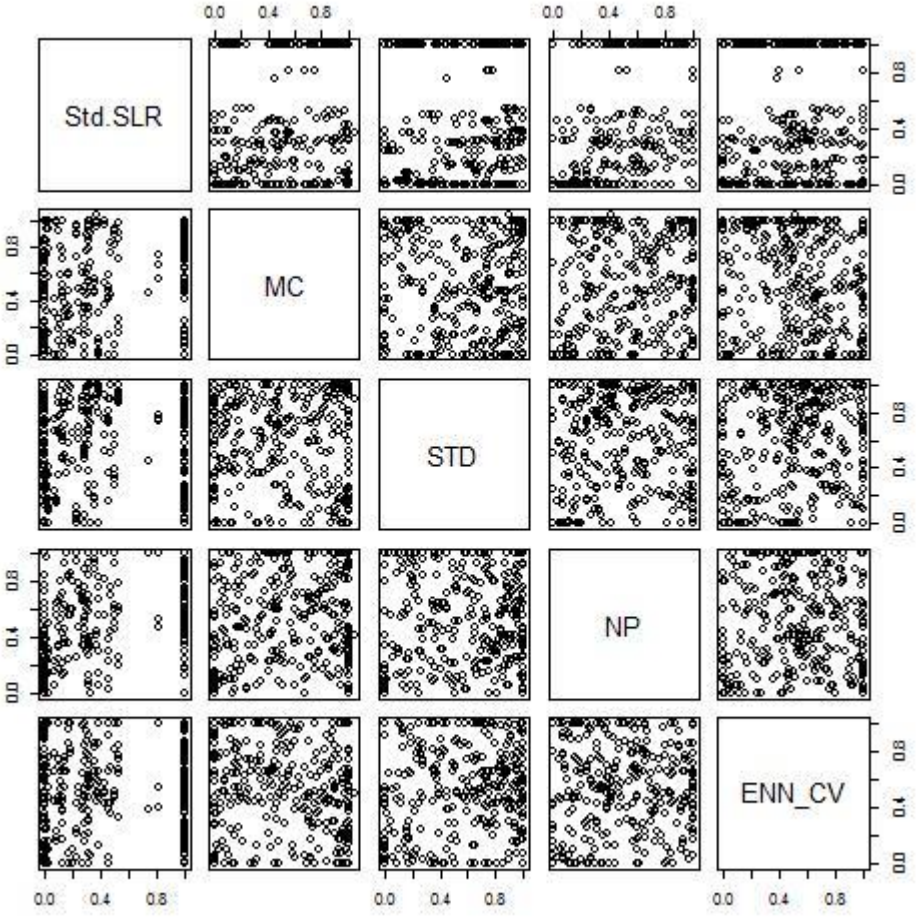
Appendix D Full list of metrics

Appendix D Full list of metrics investigated for evaluating expansion patterns

Software	Basis	Metric's full name
FRAGSTATS (McGarigal *et al.*, 2012)	Raster-density	Euclidean Nearest Neighbor (with mean, SD, and CV), Perimeter to Area ratio, Class area, Percent of landscape, Number of Patches, Total edge, Largest Patch Index, and Splitting Index
ArcGIS (Environmental Systems Research Institute, 2014)	Vector--distance	Mean centre, Standard distance, Point distance (with mean, maximum, and SD)
Manual calculation	Raster--density	Schwartzberg measure of compactness

Appendix E Associations among final metrics

Appendix E Associations among final metrics used to quantify spatiotemporal patterns in species spread. Metrics exhibited weak relationships among each other (Pearson's $r \leq 0.53$), suggesting that each communicated different aspects of spatiotemporal patterns. MC=mean centre, STD=standard distance, NP=number of patches, and ENN_CV=coefficient of variation of Euclidean nearest neighbor. Prior to analysis, we unit-base normalized the species invasion curves to facilitate cross-species comparison (see Equation 1 in text



Appendix F Highway association

Appendix F 19 species showed significant association with highway networks p ≤ .07. Six species showed negative association at probability p ≤ .10. Six species showed negative association p > .96, while four species showed no significant difference from the random null distribution.

SPECIES	MRT	OBSERVED MEAN	LOW RANDOM	MEAN RANDOM	HIGH RANDOM	P-VALUE	VARIANCE
AIAL	173	0.2365	0.2263	0.2279	0.2293	.01	0.00866
ALJU	143	0.2145	0.2031	0.2043	0.2055	.01	0.01019
ALPE4	146	0.2299	0.2210	0.2228	0.2245	.01	0.00713
CEOR7	174	0.2485	0.2345	0.2365	0.2386	.01	0.01192
EUAL8	128	0.2494	0.2353	0.2377	0.2400	.01	0.01170
LECU	86	0.2118	0.2023	0.2032	0.2041	.01	0.00867
LIVU	192	0.2498	0.2370	0.2394	0.2412	.01	0.01039
LOJA	146	0.2131	0.2063	0.2072	0.2081	.01	0.00596
LOMA6	114	0.2398	0.2295	0.2307	0.2326	.01	0.00901
LOTA	141	0.2202	0.2101	0.2119	<.001	.01	0.00829
MEAZ	129	0.1996	0.1912	0.1926	0.1940	.01	0.00704
MISI	165	0.2344	0.2256	0.2277	0.2294	.01	0.00669
MIVI	83	0.2381	0.2276	0.2292	0.2307	.01	0.00893
PUMOL	114	0.2064	0.2034	0.2046	0.2062	.01	0.00178
ROMU	118	0.2232	0.2154	0.2169	0.2181	.01	0.00626
VIMA	145	0.2205	0.2117	0.2134	0.2149	.01	0.00706
VIMI2	192	0.2442	0.2364	0.2377	0.2390	.01	0.00657
PATO2	142	0.2159	0.2132	0.2146	0.2159	.01	0.00131
WISI	135	0.2140	0.2110	0.2127	0.2148	.07	0.00130
IMCY	91	0.1622	0.1595	0.1614	0.1636	.14	0.00086
NADO	135	0.2148	0.2119	0.2141	0.2158	.15	0.00070
LOMO2	115	0.2303	0.2279	0.2303	0.2335	.47	0.00000
LISI	110	0.2010	0.1999	0.2010	0.2020	.49	0.00000
TRSE6	163	0.1708	0.1706	0.1719	0.1737	.96	-0.00117
ELAN	141	0.2156	0.2151	0.2180	0.2204	.99	-0.00240
ELUM	141	0.2192	0.2216	0.2231	0.2244	1	-0.00393
EUFO5	126	0.2222	0.2223	0.2241	0.2256	1	-0.00185
LYJA	133	0.1765	0.1772	0.1786	0.1800	1	-0.00219
ROBR	146	0.1850	0.1872	0.1899	0.1920	1	-0.00495